



universität
wien

Diplomarbeit

Titel der Arbeit

Die Untersuchung der Theory of Mind im Rahmen der
Bewältigung von EMS-ähnlichen Aufgaben in
Abhängigkeit von Stereotype Threat, Risikoverhalten
und Geschlecht

Verfasserin

Katrin Anzirk

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im Juli 2011

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: A.o.Univ. Prof. Dr. Mag. Ulrike Willinger

Danksagung:

In erster Linie möchte ich mich bei meinen Eltern und bei meinem Freund Michi, für all ihre Unterstützung, Hilfe, Kraft und ihre unendliche Geduld bedanken. Sie haben mich in den schwierigen Momenten beruhigt, ermutigt und motiviert.

Ein riesiges Dankeschön auch an Bernd O., der mir mit Rat und Tat zur Seite stand.

Mein Dank gilt auch meinen Diplomarbeitkolleginnen Gabi, Helene, Silvia und Anita für die Unterstützung, das gegenseitige Aufbauen und das gemeinsame Bewältigen dieser schwierigen Zeit.

Ein großes Dankeschön an meine betreuende Professorin, Ao. Univ. Prof. Mag. Dr. Ulrike Willinger.

Zu guter Letzt möchte ich mich auch bei allen Direktoren, Professoren und SchüleInnen bedanken, durch die diese Untersuchung überhaupt erst möglich gemacht wurde.

Anmerkung:

Die Studie, die im Rahmen dieser Diplomarbeit vorgestellt wird, wurde in Zusammenarbeit mit vier weiteren Diplomandinnen, Gabriele Hangl, Silvia Hameseder, Helene Lager und Anita Teufl, durchgeführt. Jede der Kolleginnen behandelt unterschiedliche Schwerpunkte in ihren Diplomarbeiten, wobei die verwendeten Daten und die Grundlage der Studie in den unterschiedlichen Abhandlungen ident sind und somit Überschneidungen in der Literatur bedingen können. Darüber hinaus beziehen sich alle Arbeiten auf eine gemeinsam erfasste Stichprobe. Dadurch sind Überlagerungen der einzelnen Diplomarbeiten unumgänglich und nicht als Plagiat zu sehen.

Inhaltsverzeichnis

Anmerkung:	II
Einleitung:.....	2
Theoretischer Teil	5
Kapitel 1 Einstiegstest für das Medizinstudium – EMS	5
1.1. Fakten zum EMS.....	5
1.2. Beschreibung des Aufbaus.....	6
1.3. Bisherige Studien zum EMS	7
Kapitel 2 Stereotype Threat	10
2.1. Was ist ein Vorurteil?.....	10
2.2. Was versteht man unter einem Stereotyp?.....	11
2.2.1 Geschlechtsspezifische Stereotype.....	12
2.3. Ursprung und Hintergrund.....	13
2.4. Definition und Beschreibung des Phänomens Stereotype Threat	14
2.5. Forschungsergebnisse zu Stereotype Threat	15
2.5.6 Stereotype Threat und <i>Stigmabewusstsein</i>	21
2.5.7 Stereotype Threat- negative Gedanken, Angst und Besorgnis.....	22
2.5.8 Stereotype Threat – Auswirkungen auf das Erregungsniveau	23
2.5.9 Stereotype Threat in Verbindung mit der Erwartungshaltung	24
2.5.10 Stereotype Threat und die Rolle der Hormone.....	24
Kapitel 3 Risikoverhalten	27
3.1. Historische und theoretische Hintergründe.....	28
3.2. Was sind Entscheidungen?.....	28
3.3. Wie entstehen Entscheidungen?.....	29
3.3.1 Präskriptive Entscheidungstheorie	30
3.3.2 Deskriptive Entscheidungsmodelle	30
3.3.3 Prospekt Theory (PT).....	30
3.4. Entscheidungen unter Risikobedingungen.....	34
3.5. Geschlechtsspezifische Unterschiede im Risikoverhalten	36
3.6. Erfassung des Risikoverhaltens.....	40
3.6.1 Game of Dice Task (GDT).....	42
Kapitel 4 Theory of Mind:	44
4.1. Geschichte und Hintergrund	44
4.2. Definition und Beschreibung	45

4.3.	Konzept der falschen Überzeugungen – False belief	46
4.4.	Entwicklung der ToM	48
4.4.1	Entwicklung einer fortgeschrittenen ToM	49
4.5.	Erklärungsansätze und Theorien zur ToM.....	50
4.5.1	Theorie-Theorie.....	50
4.5.2	Simulationstheorie.....	51
4.5.3	Modularitätstheorie	52
4.6.	Geschlechtsspezifische Unterschiede.....	53
4.6.1	Unterschiede im Verständnis falscher Überzeugungen.....	53
4.6.2	Unterschiede in der sozialen Kompetenz	55
4.6.3	ToM und Autismus	57
4.7.	Wechselwirkungen und Zusammenhänge.....	61
Kapitel 5	Zielsetzungen, Fragestellungen und Hypothesen.....	64
5.1.	Zielsetzungen und Fragestellungen.....	64
5.2	Hypothesen.....	65
5.2.1	Haupthypothese.....	65
5.2.2	Hypothesen zur Theory of Mind	65
5.2.4	Nebenhypothesen zum Game of Dice Task (GDT)	68
II.	Empirischer Teil	69
Kapitel 6	Methode	69
6.1	Untersuchungsplan und Intendierte Stichprobe	69
6.2.	Vorstudie.....	73
6.3.	Erhebungsinstrumente.....	74
6.3.1	Soziodemografischer Fragebogen	74
6.3.2	<i>Wissenstest</i>	74
6.3.3	EMS-Untertests	76
6.3.4	Leistungsmotivationsinventar (LMI)	79
6.3.5	Skala zur Allgemeinen Selbstwirksamkeitserwartung	80
6.3.6	ToM-Aufgaben.....	81
6.3.7	Game of Dice Task (GDT).....	83
Kapitel 7	Untersuchung:	84
7.1.	Untersuchungsdurchführung	84
7.2	Stichprobe	85
7.2.1	Geschlechterverteilung innerhalb der Gesamtstichprobe.....	85

7.2.2	Alter der SchülerInnen	91
7.2.3	Schulzweige und Staatsbürgerschaft.....	93
7.2.4	Wunsch, Medizin zu studieren und Vorbereitungen für EMS	95
Kapitel 8	Ergebnisse	99
8.1	Überprüfung der experimentellen Variation	99
8.1.1	Reliabilitätsanalysen	99
8.1.1.2	GDT	104
8.2	Auslösung der Stereotype Threat - Bedingungen	104
8.3	Deskriptive Ergebnisse der Erhebungsinstrumente.....	126
8.3.1	Ergebnisse der EMS-ähnlichen Aufgaben	127
8.3.2	Deskriptive Ergebnisse des Game of Dice Task(GDT)	137
8.3.3	Deskriptivstatistik ToM-Aufgaben	138
8.4	Ergebnisdarstellung der Hypothesen.....	139
8.4.1	Univariate Ergebnisse der ToM	139
8.2.2	Zusammenhang zwischen ToM und den EMS-ähnlichen Aufgaben	145
8.2.3	Nebenhypothesen zur ToM	148
8.2.4	Nebenhypothese zum GDT	154
Kapitel 9	Diskussion:.....	158
Kapitel 10.	Zusammenfassung.....	170
11.	Abstract	174
	Literaturverzeichnis:	175
	Tabellenverzeichnis:	185
	Abbildungsverzeichnis:.....	189
Anhang:	190
A1.	Elternbrief	190
A2.	Information für LehrerInnen	191
A3.	Einverständniserklärung.....	195
A4.	Lebenslauf	199

Einleitung:

„Bestehen oder Nicht-Bestehen?“ Eine Frage, eine Tatsache, ein Fluch? Wie oft stellen wir uns, in den unterschiedlichsten Situationen diese Frage? Kennt nicht jeder von uns, egal ob Mann oder Frau, mindestens eine Situation in seinem/ihrem Leben, in dem er/sie unter dem Druck stand, in einer gewissen Situation unter schwierigen Bedingungen, bestehen zu müssen oder zu wollen? Erinnert sich nicht jeder von uns an eine oder mehrere Situationen, in denen er/sie auf Abruf, unter schwierigen und zeitlich begrenzten Bedingungen und unter komplexen Arbeitsanforderungen eine sehr gute Leistung erbringen musste, die sie oder ihn an die Grenze seiner Fähigkeiten führte? Und kennt nicht auch jeder von uns die Aufregung, die Angst und den Druck, um der eigenen Erwartungshaltung und jener der anderen gerecht zu werden? In diesen außergewöhnlichen Situationen, die von den genannten emotionalen Zuständen begleitet werden und Auswirkungen auf den unterschiedlichsten Ebenen für die jeweilige Person haben, sind oftmals, teils bewusst und teils auch unbewusst bestimmte Klischeevorstellungen und Vorurteile präsent. Zur Veranschaulichung möchte ich ein Beispiel bringen, das auch den Bezug zu einem wichtigen Konstrukt, dem Phänomen *Stereotype Threat* oder *Bedrohung durch Stereotype* im Rahmen dieser Diplomarbeit darstellt. Es handelt sich dabei um folgende Situation: Eine junge Frau möchte ein Medizinstudium beginnen. Der Beruf des Arztes ist in unserer Gesellschaft nach wie vor noch ein sehr männlich dominiertes Arbeitsfeld, obwohl es genügend Frauen gibt, die sich für diesen Beruf interessieren und die dafür geforderten Qualifikationen besitzen. Es herrschen jedoch die Vorurteile und Stereotype vor, dass Frauen eine Schwäche in den kognitiven Bereichen des Schlussfolgernden Denkens und der Raumvorstellungskraft besitzen. Dieses Wissen und Bewusstsein über diese frauenspezifischen Vorurteile und Stereotype haben beide Geschlechter verinnerlicht. Befindet sich nun eine Frau in der Aufnahmesituation für die Zulassung für dieses Studium, so sind diese Vorurteile und Stereotype präsent und können eine gewisse Bedrohung für die Frau darstellen, da diese unweigerlich bemüht ist, die vorherrschenden Stereotype nicht zu bestätigen. Das Wissen über die Existenz dieser Stereotype und die dadurch empfundene situationsspezifische Bedrohung, sowie das Bemühen darum diese

nicht zu bestätigen wirken sich negativ und hemmend auf die Leistung der Frau im Aufnahmeverfahren aus und können dazu führen, dass die Frau, trotz ihrer prinzipiell vorhandenen Fähigkeiten die Aufnahmeprüfung nicht besteht und somit auch nicht die Ausbildung antreten kann. Diese bedeutende und komplexe Thematik der *Bedrohung durch Stereotype* (auch *Stereotype Threat* genannt) und deren Auswirkungen auf die betroffenen Personen werden im Rahmen dieser Studie aufgegriffen und im Zusammenhang mit der *Theory of mind (ToM)* und dem *Risikoverhalten* untersucht, wobei der Schwerpunkt der vorliegenden Arbeit auf der Untersuchung der *Theory of Mind* liegt. Die *Theory of Mind* ist die grundlegende und essentielle Fähigkeit, sich in andere Menschen hineinversetzen zu können und die eigenen Gedanken, Gefühle und Absichten von denen der anderen zu unterscheiden, im Bewusstsein, dass die Handlungen anderer durch ihre emotionalen und mentalen Zustände bedingt sind (Förstl, 2007) Eine gut entwickelte *ToM* ist ausschlaggebend und unerlässlich für Kernkompetenzen wie zum Beispiel soziales Verständnis, soziale Kompetenz, Empathiefähigkeit, Perspektivenübernahme, moralisches Verständnis sowie Rücksichtnahme und Respekt. Es ist das Anliegen dieser empirischen Arbeit eine Verbindung zwischen der *ToM*, dem Phänomen *Stereotype Threat* und dem Risikoverhalten unter Berücksichtigung geschlechtsspezifischer Unterschiede herzustellen. Der Neuwert dieser Studie besteht einerseits darin, dass die VersuchsteilnehmerInnen vier unterschiedlichen *Stereotype Threat*-Versuchsbedingungen zugeteilt wurden (Aktivierung, Neutral, Aufklärung und Negativ-Aktivierung), wobei die Versuchsbedingung der *Negativ-Aktivierung* („Frauen sind besser in logischen und räumlich-visuellen Aufgaben“) in der Literatur bis dato noch nicht vorgegeben wurde. Andererseits wird auch die Untersuchung der *ToM* im Zusammenhang mit dieser Thematik in dieser Form zum ersten Mal durchgeführt..

Hintergrund und Anstoß für die Planung und Realisierung dieser Studie war die in den Medien bereits viel diskutierte Thematik des *EMS*, des Eingangstests für das Medizinstudium und die damit verbundene Problematik des schlechten Abschneidens von weiblichen Anwärtern im Gegensatz zu ihren männlichen Kollegen. Eine von Spiel, Schober und Litzenberger durchgeführte Evaluierung

des EMS im Jahr 2007 konnte diesbezüglich deutliche Schwachstellen aufzeigen (siehe Kapitel 1). Infolgedessen wurden Fragen bezüglich möglicher Ursachen und Einflussfaktoren aufgeworfen. Frau Professor Mag. Dr. Willinger, unsere betreuende Professorin, erhielt von der Medizinischen Universität Wien den Auftrag in Kooperation mit der Universität Wien eine Studie durchzuführen, die diese Thematik aufgreifen sollte. Meine Diplomarbeitkolleginnen und ich wurden in weiterer Folge beauftragt soziale und biologische Einflussfaktoren auf die Leistung in Testsituationen bei Burschen und Mädchen zu untersuchen.

Im Rahmen dieser Diplomarbeit soll, wie bereits erwähnt, die *Theory of Mind* im Rahmen der Bewältigung EMS-ähnlicher Aufgaben in Abhängigkeit von Stereotype Threat, dem Risikoverhalten und dem Geschlecht unter Berücksichtigung geschlechtsspezifischer Unterschiede untersucht werden.

Die Arbeit gliedert sich in folgende Teile: Der theoretische Teil dieser Diplomarbeit befasst sich mit der eingehenden Darstellung und Erklärung der drei zu untersuchenden Konstrukte: *Stereotype Threat*, *Risikoverhalten* und *ToM*. Diesbezüglich wird der bisherige Stand der Forschung in Bezug auf die unterschiedlichen Themenbereiche aufgearbeitet-

Der empirische Teil hingegen widmet sich der Beschreibung der Untersuchungsdurchführung und der Erhebungsinstrumente, der Darstellung, Erklärung und Auswertung der Fragestellung sowie abschließend der Interpretation und Diskussion der Ergebnisse.

Theoretischer Teil

Kapitel 1 Einstiegstest für das Medizinstudium – EMS

Im folgenden Kapitel wird der EMS, der Einstiegstest für das Medizinstudium im Allgemeinen kurz vorgestellt. Es werden sowohl wichtige Eckdaten und Fakten angeführt, als auch kurz auf den Aufbau des Tests und bisherige Studien und Forschungsergebnisse eingegangen.

1.1. Fakten zum EMS

„Aufgrund des Urteils des Europäischen Gerichtshofs vom 7.7.2005, mit dem die bisherige Regelung über den Zugang zu österreichischen Universitäten als europarechtswidrig qualifiziert wurde, und der damit geschaffenen neuen Rechtslage, die zu einem verstärkten Andrang von Studierenden aus dem Ausland, insbesondere aus Deutschland, geführt hat, werden die Medizinischen Universitäten in Wien und Innsbruck auf Basis der vom Nationalrat erlassenen Novelle zum Universitätsgesetz 2002 gemeinsam eine kapazitätsorientierte Studienplatzvergabe für ALLE StudienwerberInnen durchführen.“([http://www.eignungstest-medizin.at/cms/\(25.04.2011\)](http://www.eignungstest-medizin.at/cms/(25.04.2011))).

„Die Vergabe der Plätze erfolgt mittels eines erprobten und wissenschaftlich abgesicherten Eignungstests (EMS), der in Deutschland entwickelt, in der Schweiz weiterentwickelt und seit 1998 angewendet wird. Der Eignungstest liefert einen Testwert, welcher nachweislich hoch mit der Studieneignung korreliert. Dieser wird für die Zulassung verwendet und bildet das Zulassungskriterium.“([http://www.eignungstest-medizin.at/cms/\(25.04.2011\)](http://www.eignungstest-medizin.at/cms/(25.04.2011))).

Der Eignungstest für das Medizinstudium (EMS) wird in der Schweiz bereits seit 1998 eingesetzt und seit 2006, gemeinsam an den Universitäten Basel, Bern, Fribourg, Zürich, den Medizinischen Universitäten Innsbruck, Wien durchgeführt. Die Teststruktur des EMS wurde 2004 in der Schweiz, infolge einer Anforderungsanalyse der neuen Studienbedingungen, erstmals revidiert und um

den Untertest „Planen und Organisieren“ erweitert (<http://www.unifr.ch/ztd/ems/EMSAufbau.pdf>, 25.04.11).

1.2. Beschreibung des Aufbaus

Der EMS basiert ursprünglich auf dem, ab 1986 in Deutschland eingesetzten, Test für Medizinische Studiengänge (TMS). Das Ziel des Tests ist die Vorhersage des Studienerfolges (Hänsgen & Spicher, 2007).

Seit 2005 besteht der EMS aus 10 Untertests (siehe Tabelle 1) mit insgesamt 198 Aufgaben, die in Abstimmung auf die Studienanforderungen konstruiert wurden und abschließend zu einem Gesamtwert verrechnet werden. Für die Bearbeitung der einzelnen Untertests gibt es jeweils eine gewisse Zeitvorgabe, die die erlaubte Bearbeitungszeit limitiert. Prinzipiell sind die Zeitvorgaben so konzipiert, dass das vollständige Bearbeiten und Beantworten aller Aufgaben unmöglich gemacht wird (<http://www.ems-eignungstest.ch/at-testaufbau.htm>, 24.04.11).

Tabelle 1: 10 EMS-Untertests und die dadurch erfassten Fähigkeiten

Untertests	Was wird damit gemessen?
Quantitative und formale Probleme	Umgang mit Zahlen und Formeln im Rahmen medizinisch-naturwissenschaftlicher Fragen
Schlauchfiguren	Räumliches Vorstellungsvermögen
Textverständnis	Verarbeitung von komplexem Textmaterial
Planen und Organisieren	Effiziente Selbstorganisation im Studium
Konzentriertes und sorgfältiges Arbeiten	Rasches und konzentriertes Arbeiten
Figuren lernen	Wahrnehmen und Einprägen von visuell dargebotenen Einzelheiten
Fakten lernen	Einprägen von Fakten
Medizinisch-naturwissenschaftliches Grundverständnis	Verständnis für Fragen der Medizin und der Naturwissenschaften

Muster zuordnen	Ausschnitte in einem komplexen Bild wieder erkennen
Diagramme und Tabellen	Analyse und Interpretation von Diagrammen und Tabellen

Die Gesamtdauer des Tests beträgt inklusive einer einstündigen Pause 8 Stunden. Laut Autoren beträgt die reine Bearbeitungsdauer des EMS 5 Stunden und 15 Minuten.

Für Graz wurde ein eigener Wissenstest konzipiert, der aus insgesamt 200 Aufgaben aus den Bereichen Biologie, Chemie, Physik, Mathematik und Textverständnis besteht. Auch hier wird das Ziel der Vorhersage der Studieneignung und ausdrücklich nicht die Vorhersage der Berufseignung verfolgt.

Tatsache ist, dass in beiden Auswahlverfahren Frauen im Mittel schlechter abschnitten als Männer und österreichische BewerberInnen schlechter als deutsche. Die Leistungsunterschiede hatten weitreichende Konsequenzen für die Geschlechts- und Nationalitätenverteilung bei den Aufgenommenen und daher wurde eine „Safeguard-Klausel“ zum Schutz der Homogenität des Bildungssystems in das UG 2002 aufgenommen und Aufnahmequoten (75% ÖsterreicherInnen, 20% EU BürgerInnen, 5% Nicht-EU BürgerInnen) für das Medizinstudium festgelegt, um damit den Anteil der aufgenommenen deutschen BewerberInnen zu limitieren. (<http://ug.manz.at>, 30.08.2011)

1.3. Bisherige Studien zum EMS

Es ist eine nicht zu leugnende Tatsache, dass es geschlechtsspezifische Unterschiede im EMS gibt. In der Vergangenheit wurde diesbezüglich häufig die Argumentation ins Treffen geführt, dass Frauen vor der Einführung des Eignungstests schlechtere Prüfungsleistungen bzw. höhere Drop-out Raten im Studium aufwiesen als Männer. Aktuellere Untersuchungen (Frischenschlager, Mitterauer & Haidinger, 2005; 2007) zeigen jedoch, dass das Geschlecht bei

eingehender und differenzierter Betrachtung kein adäquater und valider Prädiktor für den Studienerfolg ist. Es ist anzumerken, dass Frauen zwar zu Studienbeginn einen niedrigeren Prüfungserfolg als Männer aufweisen, sich diese anfänglichen Unterschiede aber im Verlauf des ersten Studienjahrs amortisieren und zu Studienende de facto mehr Frauen das Medizinstudium erfolgreich absolvieren als Männer (Frischenschlager, Mitterauer & Haidinger, 2005, 2007).

Frischenschlager, Mitterauer und Haidinger (2005) beschäftigten sich im Rahmen ihrer Untersuchung mit der Fragestellung, ob Leistungsfaktoren wie Schulnoten, fachspezifische Vorkenntnisse und sowohl die allgemeine als auch die spezifische Lernleistung potentielle Auswahlkriterien und Prädiktoren für das Medizinstudium und den Studienerfolg darstellen. Die Ergebnisse zeigen, dass es besonders bei den Schulleistungen, den fachspezifischen Vorkenntnissen und auch bei der studienspezifischen Lernleistung signifikante Unterschiede zwischen Erfolgreichen und Nicht-erfolgreichen gibt. Jedoch weist keiner der untersuchten Faktoren eine ausreichende Trennschärfe auf und kann somit auch nicht als aussagekräftiges und valides Auswahlkriterium herangezogen werden.

Des Weiteren haben sich Spiel, Schober und Litzenberger (2007) mit dieser Thematik auseinandergesetzt und wurden im Juni 2007 dazu beauftragt, im Vorfeld des Auswahlverfahrens für das Studienjahr 2007/2008, den Eignungstest für das österreichische Medizinstudium zu evaluieren. Das Hauptaugenmerk lag auf der Untersuchung und Überprüfung der Fairness des Eignungstests in Bezug auf das Geschlecht, die Nationalität, die soziale Herkunft (Schulbildung der Eltern) und den Schultyp (naturwissenschaftlich versus nicht-naturwissenschaftlich). Die Evaluationsstudie von Spiel et al. (2007) wurde als Fragebogenerhebung, die im Anschluss an den EMS-Test stattfand, durchgeführt. Der Fragebogen wurde größtenteils von den AutorInnen selbst konstruiert und bestand aus 36 Fragen, die in einer begrenzten Zeit von fünf Minuten vorgegeben wurden. Es wurden zu den Bereichen familiärer Hintergrund/Herkunft, schulischer Hintergrund, Erfahrungen und Umgang mit dem Medizinauswahlverfahren, Selbsteinschätzungenbezogen auf das Medizinauswahlverfahren und Relevanz des Medizinstudiums

Erklärungsvariablen erhoben, die für die Aufklärung der Leistung im Test herangezogen wurden (Spiel et al., 2007).

Im Zuge der Evaluierung konnten signifikante geschlechtsspezifische Leistungsunterschiede zu Gunsten der Männer festgestellt werden (siehe auch Mallinger et al., 2009). Dieser Unterschied war besonders deutlich in den drei Untertests *Quantitative und formale Probleme*, *Schlauchfiguren* und *Diagramme und Tabellen*. In Bezug auf die Trennschärfe der Aufgaben und der Reliabilität der Untertests sind die Ergebnisse weniger zufriedenstellend, da die meisten Aufgaben sehr geringe Trennschärfen aufweisen, was bedeutet, dass diese Aufgaben nicht gut geeignet dafür sind, um zwischen „guten“ und „schlechten“ Testpersonen zu unterscheiden. Weiteres merken die Autoren an, dass die einzelnen Untertests über geringe Reliabilitäten verfügen, was den Grad der Genauigkeit angibt, mit der ein Untertest ein bestimmtes Merkmal misst. Im Zuge der Analysen zeigten sich ebenfalls Probleme mit der Verrechnungsfairness.

Des Weiteren konnte festgestellt werden, dass die Schulbildung der Eltern eine relevante Einflussgröße darstellt, denn umso höher die Schulbildung der Eltern ist, desto besser sind die Testleistungen bei gleichen Schulnoten.

Zusammenfassend ist zu sagen, dass im Zuge der Evaluierung des EMS mehrere Schwachstellen und Probleme identifiziert werden konnten, die sich nachteilig auf die Aufnahme und Zulassung von weiblichen Anwärtern auswirken. Ein großer und nennenswerter Kritikpunkt ist in diesem Zusammenhang die fehlende geschlechtsspezifische Normierung.

Kapitel 2 Stereotype Threat

In unserer Gesellschaft bestehen zahlreiche Vorurteile, die das Weltbild, die Weltanschauung, die Haltung und Einstellung gegenüber Menschen und Situationen prägen. Diese Vorurteile und Stereotype werden oftmals herangezogen, um Handlungen und Eigenschaften von Menschen, Dingen, Vorkommnissen begreifbar und nachvollziehbar zu machen. Sie verleiten jedoch zu einer voreiligen und sehr subjektiven Meinungsbildung und verursachen *Schwarz-Weiß-Denken* und *Etikettierung* (Becker, 1981). Stereotype beschreiben immer eine bestimmte Gruppe von Individuen und sind sehr weit verbreitet, sowohl bei der betroffenen Gruppe selbst als auch bei den anderen Menschen.

Vorurteile und Stereotype findet man in den unterschiedlichsten Bereichen. In der Literatur beschäftigte und untersuchte man bisher insbesondere 3 Themengebiete in Bezug auf die Wirkung von Stereotypen und dem Phänomen *Stereotype Threat*:

- Geschlechtsspezifische Unterschiede und Geschlechtsstereotype: "Frauen sind schlechter in mathematischen und naturwissenschaftlichen Bereichen als Männer" (Spencer, Steele & Quinn, 1998; Cadinu et al., 2005),
- Kulturelle und ethnische Unterschiede und Stereotype: "Afrikaner erbringen schlechtere akademische Leistungen als Weiße" (Steele & Aronson, 1995, Aronson et al., 1998; Aronson et al., 1999),
- Altersspezifische Unterschiede und Stereotype: „ältere Menschen sind vergesslicher“ (Levy & Langer, 1994).

2.1. Was ist ein Vorurteil?

In der Sozialpsychologie versteht man unter dem Begriff Vorurteil eine negative, abwertende und feindselige Einstellung gegenüber Personen, die einer bestimmten Gruppe angehören. Diese negative Einstellung gegenüber den Mitgliedern der jeweiligen Gruppe wird alleinig aufgrund der Tatsache der Zugehörigkeit zu dieser Gruppe gebildet (Aronson, Wilson & Akert, 2008). Vorurteile stehen somit für eine bestimmte, meist negative Einstellung und die damit verbundene emotionale Komponente. Bei Vorurteilen handelt es sich um negative und

verallgemeinernde Merkmalszuschreibungen, die sich auf die gesamte Gruppe beziehen (z.B. „Schwarze sind dumm“, „Frauen können nicht einparken“, „Arbeitslose sind Schmarotzer“) und wonach ein Mitglied dieser Gruppe bewertet wird. Individuelle Eigenschaften und Unterschiede werden nicht berücksichtigt und weitgehend ignoriert. Personen, die Vorurteile gegenüber einer bestimmten Gruppe hegen, verhalten sich abweisend, abwertend und distanziert. Vorurteile betreffen somit die emotionale, kognitive und verhaltensbezogene Ebene, da sie eine negative und abwertende, auf unsere Einstellungen und Emotionen bezogene Reaktion verursachen und darstellen (Aronson et al., 2008).

2.2. Was versteht man unter einem Stereotyp?

Die Bezeichnung des Stereotyps geht zurück auf den Journalisten Walter Lippmann und fand 1922 Eingang in die Sozialwissenschaften. Lippmann versteht darunter bestimmte vorgefertigte *Bilder*, die wir von verschiedenen Menschen und Gruppen haben und in uns tragen. Diese Bilder sind subjektiv und stehen im Kontrast zu der Realität (Aronson et al., 2008).

Stereotype beinhalten im Gegensatz zu Vorurteilen sowohl positive als auch negative Merkmalszuschreibungen. Laut Aronson et al., (2008, S.425) ist ein Stereotyp „eine verallgemeinernde Annahme über ein Gruppe von Menschen, die praktisch all ihren Mitgliedern, unabhängig von tatsächlichen Unterschieden zwischen ihnen, bestimmte Eigenschaften zuschreibt“.

Allgemein werden Stereotype zur Kategorisierung von Menschen eingesetzt und sind veränderungsresistent. Bei der Stereotypisierung handelt es sich um einen kognitiven Prozess, der uns dabei hilft, Menschen schnell und oberflächlich einzuschätzen und die Welt, in für uns vereinfachter Form darzustellen und zu begreifen. (Aronson et al., 2008).

Eine weitere Erklärung und Beschreibung der Stereotypisierung liefert Allport (1954) mit dem *Gesetz der geringsten Anstrengung*. Demnach entwickeln wir, einerseits aufgrund der nicht zur Gänze bewältigbaren Informationsflut aus der Umwelt und andererseits aufgrund der zeitlich begrenzten Kapazitäten, nur zu

gewissen Themenbereichen explizite und genaue Einstellungen und zu anderen Themen hingegen, vereinfachte, oberflächliche und generalisierte Meinungen. (Aronson et al., 2008). Das Zurückgreifen auf diese *Faustregeln* und *Abkürzungen* ist ein Versuch andere zu verstehen und rasch einschätzen zu können (Fiske & Depret, 1996)

2.2.1 Geschlechtsspezifische Stereotype

In unserer Gesellschaft sind Stereotype, die sich auf geschlechtsspezifische Unterschiede beziehen, sehr präsent und tief verankert. Ein Beispiel dafür wäre, dass Frauen generell als fürsorglicher und weniger durchsetzungsfähig gelten als Männer (Aronson et al., 2008). Ob dieses Stereotyp allein auf der Wahrnehmung traditioneller Rollenbilder oder auf biologischen Faktoren basiert, ist wissenschaftlich nicht eindeutig geklärt. Tatsache ist, dass Geschlechtsstereotype weit verbreitet und wie bereits erwähnt, tief verankert im gesellschaftlichen Denken und der Wahrnehmung sind. Diese Geschlechtsstereotype sind oftmals nicht konsistent mit der Realität und können daher für die betroffenen Personen erhebliche Auswirkungen haben. Zu ähnlichen Erkenntnissen kamen Forscher bereits vor über dreißig Jahren wie sich in Experimenten von Feldman-Summers und Kiesler (1974) und Deaux und Emsweiler (1974) zeigt. In einem Experiment von Feldman-Summers und Kiesler (1974) schätzten männliche Studierende eine erfolgreiche Ärztin als weniger kompetent ein als einen erfolgreichen Arzt, da sie der Meinung waren, dass es die Frau leichter gehabt hätte, Karriere zu machen. Die weiblichen Studentinnen hingegen schätzten beide als gleich kompetent ein und meinten, dass es der Arzt leichter gehabt hätte.

Deaux und Emsweiler (1974) fragten weibliche und männliche Studenten im Rahmen ihrer Untersuchung, worauf die gute Leistung eines/r Mitstudenten/in in einer schwierigen Aufgabe zurückzuführen sei. Es konnte demonstriert werden, dass im Falle eines männlichen Mitstudierenden die gute Leistung von beiden Geschlechtern auf seine Fähigkeit zurückgeführt wurde. Im Unterschied dazu meinten beide Geschlechter im Falle einer Frau, dass ihre Leistung auf Glück

basieren würde. Damit konnte einerseits gezeigt werden wie manifest geschlechtsspezifische Stereotype sind und andererseits, dass diese sexuellen Stereotype so stark sein können, dass sie von Mitgliedern der betroffenen Gruppe sogar selbst übernommen werden (Deaux & Emsweiler, 1974).

Diese Ergebnisse konnten auch in weiterführenden Studien 30 Jahre später repliziert werden und veranschaulichen, dass geschlechtsspezifische Stereotype negative Auswirkungen für die betroffenen Personen, in diesem Fall für Frauen, haben können.

In diesem Zusammenhang ist die Bedrohung durch negative Stereotype anzuführen. Wie bereits erwähnt, können Stereotype so stark und präsent sein, dass sie sogar von den betroffenen Mitgliedern einer bestimmten Gruppe übernommen und verinnerlicht werden. Dieses Wissen und Bewusstsein über negative Stereotype über die eigene Gruppe kann in bestimmten Situationen, in denen Stereotype aktiviert werden, dazu führen, dass bei der betroffenen Person im Zuge dieser Bedrohung durch negative Stereotype Angst ausgelöst wird. Dieses Phänomen wird *Bedrohung durch Stereotype* oder *Stereotype Threat* bezeichnet und konnte in einigen Studien von Steele (1997) und Steele und Aronson (1995) eindrucksvoll demonstriert werden.

Im Folgenden wird nun dieses Phänomen genauer beschrieben und erklärt.

2.3. Ursprung und Hintergrund

Der Ursprung der wissenschaftlichen Untersuchung und Auseinandersetzung des Stereotype Threat liegt in der Sozialpsychologie. Das Konzept und die Theorie der Stereotype Threat wurden von dem Sozialpsychologen Claude Steele und seinen Kollegen entwickelt. In den Anfängen der Erforschung des Phänomens Stereotype Threat publizierten Steele und Aronson (1995) mit *Stereotype Threat and the intellectual test performance of African Americans* und Steele (1997) mit *A threat in the air: How stereotypes shape the intellectual identities and performance* zwei wegweisende Studien in diesem Bereich. In diesen Studien wurde Stereotype Threat erstmals im Zusammenhang mit den verbreiteten

Stereotypen und Vorurteilen über Afro-Amerikaner in Bezug auf ihre schlechteren akademischen Leistungen untersucht. Es zeigte sich, dass Afro-Amerikaner, die vor der Testung einerseits die Instruktion, dass bei dem folgenden Test ihre intellektuellen Fähigkeiten geprüft werden würden und andererseits die Stereotype Threat aktivierende Information erhielten, dass Afro-Amerikaner anderen ethnischen Gruppen intellektuell unterlegen sind, deutlich schlechtere Leistungen erbrachten als die Kontrollgruppe, die keine Stereotype Threat Instruktion erhielt. Steele (1997) beschäftigte sich insbesondere mit den negativen Effekten, die Stereotype Threat auf die Personen der betroffenen, stigmatisierten Gruppe hat. Er versuchte zu verstehen und zu erforschen, wie und in welchem Ausmaß Stereotype Threat sowohl die Leistungen von Afro-Amerikaner bei akademischen Tests als auch jene von Frauen bei mathematisch-naturwissenschaftlichen Tests beeinflusst. In seinen Überlegungen ging Steele davon aus, dass im Zusammenhang mit dem Phänomen des Stereotype Threat nicht biologisch bedingte Gruppenunterschiede, sondern die von der Person erfahrene Stigmatisierung eine Rolle spielt. Im Rahmen der Dissertation seines Studenten Steven Spencer 1999, führte Steele eine Untersuchung durch, in der gezeigt werden konnte, dass Frauen, bei denen vor der Durchführung des mathematischen Tests, mittels einer, ihre soziale Identität bedrohenden, genderspezifischen Instruktion, Stereotype Threat aktiviert wurde, wesentlich schlechtere Leistungen erbrachten als jene Frauen, die keine derartige Instruktion erhielten. Infolge der Ergebnisse dieser Untersuchung fand das Phänomen des Stereotype Threat 1999 endgültig seinen Eingang in die Literatur und wurde infolgedessen zum Gegenstand zahlreicher wissenschaftlicher Untersuchungen (Aronson et al. 2002; Marx & Stapel, 2006).

2.4. Definition und Beschreibung des Phänomens Stereotype Threat

Laut Steele (1997) beschreibt der Begriff des *Stereotype Threat* ein situationsspezifisches Phänomen, bei dem negative Stereotype, die einer bestimmten Gruppe zugeschrieben werden, aktiviert werden und dadurch die eigenen kognitiven Fähigkeiten negativ beeinflussen können. Es entsteht dadurch,

dass die betroffenen Personen in speziellen Situationen, wie zum Beispiel in einer Testsituation, Angst haben, die über ihre Gruppe bestehenden Vorurteile und Stereotype zu bestätigen und danach beurteilt zu werden. Diese empfundene situationsbedingte Bedrohung und daraus resultierende Besorgnis und Angst führen im Sinne der *selbsterfüllenden Prophezeiung* zu einer tatsächlichen Bestätigung des Vorurteils bzw. des Stereotyps. Personen, die das Ziel solcher Vorurteile darstellen, sind sich dessen auch selbst bewusst. (Spencer, Steele & Quinn, 1998).

Das Phänomen des Stereotype Threat tritt in kritischen Situationen auf, insbesondere wenn sich die Vorurteile auf bestimmte Fähigkeiten beziehen, die in einer speziellen Situation (z.B. absolvieren eines Mathematiktests) eingesetzt, gezeigt und erfasst werden sollen. Eine Testsituation, in der diese Fähigkeit und Leistung erbracht werden soll, löst somit zusätzlichen Druck bei den betroffenen Personen aus und aktiviert einerseits das Wissen um die vorherrschenden Vorurteile und Stereotype, die über die eigene Gruppe bestehen und andererseits eine Bedrohung des eigenen Selbst, der eigenen Identität. Diese Beschäftigung und Auseinandersetzung mit der situationsspezifischen Bedrohung führt zu einer Beeinträchtigung der geforderten Leistung. Daraus resultiert in den meisten Fällen ein schlechteres Abschneiden in Testsituationen (Spencer, Steele & Quinn, 1998).

Die Stereotype Threat Theorie besagt, dass die Testsituation, die Hinweisreize auf negative Stereotype enthält, zusätzlich belastend wirkt und zu einer erheblichen Verschlechterung der Leistung in diesem Bereich führt (Steele & Aronson, 1995; Steele, 1997; Spencer et al., 1999).

2.5. Forschungsergebnisse zu Stereotype Threat

Im folgenden Teil findet eine kurze und prägnante Auseinandersetzung und Aufarbeitung der bestehenden Literatur in Bezug auf das Phänomen Stereotype Threat statt. Es werden ausgewählte Studien präsentiert, deren Hauptaugenmerk einerseits auf der Aktivierung von Stereotype Threat bei Frauen und der daraus resultierenden schlechteren Leistung in Testsituationen im Vergleich zu Männern liegt, als auch auf der Minimierung der empfundenen Bedrohung durch

unterschiedliche Manipulationen der Versuchsbedingungen. Des Weiteren werden mögliche Mediatorvariablen des Stereotype Threat angeführt, wie zum Beispiel Erfolgserwartung, Angst, Selbstwirksamkeitserwartung, Problemlösestrategien, die ebenfalls Auswirkungen auf die Leistung haben.

2.5.1 Geschlechtsspezifische Unterschiede

In der Literatur wurden bereits zahlreiche Studien durchgeführt, die sich mit der geschlechtsspezifischen Problematik bei mathematischen Testleistungen befassen und zeigen, dass Frauen in mathematischen Tests, die einen höheren Schwierigkeitsgrad aufweisen, schlechter abschneiden als ihre männlichen Kollegen. Bei mathematischen Tests mit geringerer Schwierigkeit hingegen schneiden Frauen genauso gut ab wie Männer (Spencer, Steele & Quinn, 1999). Infolgedessen suchte man nach Erklärungen für die schlechten Testergebnisse der Frauen.

In unserer männerdominierten Welt wird Frauen vermittelt, dass sie über schlechtere mathematische Fähigkeiten und Kenntnisse verfügen als Männer. Frauen haben diese Vorurteile verinnerlicht und reagieren oftmals bewusst, aber in vielen Situationen auch ganz unbewusst auf diese Stereotype. Die Angst diese Vorurteile zu bestätigen und danach beurteilt und bewertet zu werden verursacht Druck, der sich negativ auf die Arbeitsleistung auswirkt (Steele, 1995; 1997; Spencer et al., 1999). Im Rahmen einer Untersuchung konnten Spencer et al. (1999) aufzeigen, dass Frauen genau in den Testsituationen schlechter abschnitten als Männer, in denen durch eine genderspezifische Instruktion und Anführung der vorherrschenden Stereotype bezüglich der mathematischen Fähigkeiten und Fertigkeiten von Frauen Stereotype Threat provoziert und aktiviert wurde. Es konnte nachgewiesen werden, dass durch die Betonung des Genderaspekts („Frauen erbringen schlechtere mathematische Leistungen als Männer“) vor der Testsituation, Stereotype Threat bei den Frauen ausgelöst wurde und dass dessen Einfluss negative Auswirkungen auf die Testleistung hatte. Des Weiteren konnte festgestellt werden, dass die Reduzierung des Stereotype Threat durch die

Aufklärung, dass bei diesem Test keine geschlechtsspezifischen Unterschiede vorhanden sind, zu einer Verbesserung der Testleistung führte.

2.5.2 Stereotype Threat – Auswirkungen auf Problemlösestrategien

Weitere Auswirkungen, die mit der Bedrohung durch Stereotype im Zusammenhang stehen wurden von Carr und Steele (2009) untersucht. Sie kamen zu der Erkenntnis, dass Stereotype Threat das starre Beharren auf und Verharren in Bearbeitungs- und Lösungsstrategien erhöht, die vormals zielführend waren, aber jetzt ineffizient und inkorrekt sind. Carr und Steele (2009) konnten im Zuge ihrer Untersuchung zeigen, dass Frauen, die in der Testsituation durch negative Stereotype bedroht wurden, nicht nur schlechter abschnitten, sondern auch kontraproduktive und ineffizientere Problemlösestrategien anwendeten. Dem zu Grunde liegt die Annahme, dass negative Stereotype die Fähigkeit und das Potenzial der betroffenen Person, sich an neue Situationen anzupassen und effizientere Strategien zu adaptieren, beeinträchtigen. Ein Beispiel zur Veranschaulichung wäre: ein Aufnahmeverfahren besteht meist aus unterschiedlichen Aufgabenblöcken, die verschiedene Bearbeitungsstrategien erfordern, um zu der richtigen Lösung zu kommen. Die Strategie, die bei der vorherigen Aufgabe zum Erfolg geführt hat, gilt nicht für die folgende Aufgabe. wie z.B.: Fortführung logischer Reihen, die ersten Reihen sind aufsteigend in zweier Schritten zu lösen, dann ändert sich jedoch die Regel und um auf die richtige Zahl zu kommen müssen die vorherigen beiden Zahlen addiert werden. Die Testperson muss innerhalb kurzer Zeit die neue Lösungsstrategie finden. Frauen neigen dazu in diesen Situationen, insbesondere unter Stereotype Threat, länger in ihrer vormals erfolgreichen, aber mittlerweile ineffizienten Lösungs- und Bearbeitungsstrategie zu verharren und in der Adaption neuer Wege sowie in der Flexibilität im Denken gehemmt zu werden.

In Testsituationen, in denen Stereotype Threat aktiviert ist, werden die Aufmerksamkeit und der Fokus der Person unweigerlich auf die Abwertung der eigenen Identität gerichtet (Carr & Steele, 2009). In diesem Zusammenhang ist zu

erwähnen, dass Stereotype Threat sowohl Besorgnis und Angst (Bosson, Haymovitz & Pinel, 2004; Spencer et al.1999,) als auch das Erregungsniveau erhöhen (Ben-Zeev et al., 2005). Dadurch werden Hinweisreize aus der Umgebung, die für die Problem- und Aufgabenbearbeitung relevant sind, vernachlässigt und unzureichend bis gar nicht wahrgenommen. Dies führt dazu, dass die betroffene Person infolge des durch Stereotype Threat erhöhten Erregungsniveaus und der gesteigerten Angst in ihrer Anpassungsfähigkeit an neue und komplexere Situationen und im Entwickeln und Adaptieren neuer und zielführender Strategien gehemmt wird.

2.5.3 Stereotype Threat - Erhöhung des prevention focus

Des Weiteren ist anzuführen, dass verschiedene Studien auch den *Primingeffekt* untersuchten. Darunter wird ein „Behaltensvorteil durch Vorerfahrungen mit dem Gedächtnismaterial“ (Zimbardo & Gerrig, 2003, S.252) verstanden. Dabei stellten sie fest, dass sowohl *Priming* mit stigmatisierten Identitäten und Persönlichkeiten (Oyserman, Uskul, Yoder, Nesse & Williams, 2007), als auch die Erfahrung von Stereotype Threat, bei Personen, einen „vorbeugenden, bewahrenden und vorsichtigeren Fokus“, im Englischen als *prevention focus* bezeichnet, hervorruft. Laut Higgins (1997) handelt es sich beim *prevention focus* um einen Zustand, in dem sich die Person Gedanken und Sorgen über ihre Verantwortung und Sicherheit macht. Dieser besagte *prevention focus* tritt in spezifischen Situationen auf, in denen die Person ein akutes Sicherheitsbedürfnis verspürt. Infolgedessen wird bei der betroffenen Person ein risiko-aversives- und vermeidendes Verhalten ausgelöst und exploratives, erkundendes und erforschendes Verhalten gehemmt. Der Fokus liegt auf dem Vermeiden von Fehlern. Die Ergebnisse von Carr und Steele (2009) veranschaulichen, dass Stereotype Threat zu einem starren Verharren in ineffizienten Lösungsstrategien führt. Diese Inflexibilität im Denken und in der Anwendung und Adaption innovativer und effizienter Verarbeitungs- und Bewältigungsstrategien könnte unter anderem ein Erklärungsversuch für die schlechte Leistung sein.

2.5.4 **Stereotype Threat – Bereichsidentifikation und soziale Identität**

Forschungsergebnisse belegen, dass geschlechtsspezifische Unterschiede, genauso wie ethnische und rassenbezogene Unterschiede, in Testsituationen durch das Phänomen Stereotype Threat ausgelöst, begünstigt und betont werden (Aronson et al. 1998; Aronson et al., 2002; Brown & Pinel, 2003). Ein wichtiger Aspekt in diesem Zusammenhang ist die von Steele (1997) bereits untersuchte und thematisierte *Identifikation mit einer bestimmten Domäne* oder *Bereichsidentifikation*, die besagt, dass Personen vulnerabler für die negativen Auswirkungen von Stereotype Threat sind, wenn sie sich mit dem jeweiligen Bereich identifizieren und dieser auch eine gewisse persönliche Wichtigkeit hat. Damit in Zusammenhang steht der Aspekt der bereits erwähnten sozialen Identität, worunter man die eigene Identität als Mitglied einer bestimmten sozialen Gruppe versteht. Der Mensch verfügt über mehrere soziale Identitäten, die in unterschiedlichen Situationen aktiviert werden. Die *Theorie der sozialen Identität* besagt, dass Menschen danach streben ein positives Bild ihrer sozialen Identität zu repräsentieren und diesem auch gerecht zu werden (Tajfel & Turner, 1986). Wird diese soziale Identität durch negative Stereotype über die eigene Gruppe bedroht, so leidet nachweislich die erbrachte Testleistung. Des Weiteren zeichnete sich ab, dass mit der Zunahme der Bedeutung einer bestimmten sozialen Identität für eine Person die Verschlechterung der Leistungen zunahm. Schmader (2002) befasste sich in seiner Studie schwerpunktmäßig mit der Bedeutung der Geschlechterrolle und der sozialen Identität im Zusammenhang mit den negativen Effekten der Stereotype Threat auf die erbrachte Leistung. Er konnte zeigen, dass Frauen, für die ihre Geschlechterrolle und soziale Identität einen sehr großen Stellenwert hatte, unter der Stereotype Threat Bedingung wesentlich schlechtere Leistungen erbrachten. Hingegen zeigte sich auch, dass Frauen, die der Geschlechterrolle nicht so viel Wichtigkeit beimessen, genauso gut im Test abschnitten wie ihre männlichen Kollegen. Diese Ergebnisse erläutern, dass die Identifizierung mit der eigenen sozialen Gruppe und deren Wichtigkeit, im

Rahmen der Selbstdefinition ausschlaggebend sind und sowohl das Auftreten als auch die Auswirkungen von Stereotype Threat beeinflussen.

2.5.5 Stereotype Threat und multiple soziale Identitäten

Rydell, McConnell und Beilock (2009) griffen in ihrer Untersuchung ebenfalls die Thematik der sozialen Identität auf. Sie bezogen in ihre Überlegungen die Annahme mit ein, dass der Mensch mehrere soziale Identitäten besitzt, die unter verschiedenen Umständen und zu situationsspezifischen Zeitpunkten präsent sind. Gemäß der *sozialen Identitätstheorie* von Taifel und Turner (1986) streben die Menschen nach einem positiven Selbstbild und Selbstkonzept. In Bezug auf die Annahme der multiplen sozialen Identitäten einer Person bedeutet dies, dass in einer Situation, in der zwei soziale Identitäten passend und zutreffend sind, jene verstärkt und zugänglich wird, die ein positives Bild der Person präsentiert und eine steigernde Wirkung auf das Selbstbewusstsein hat. Gleichzeitig wird der Zugang und Abruf der mit negativen Stereotypen behafteten Identität verhindert (Rydell et al.2009).

Schmader et al. (2008) haben ein Prozessmodell zur Erklärung und Veranschaulichung von Stereotype Threat im Zusammenhang mit der sozialen Identität und den Auswirkungen auf das Arbeitsgedächtnis entwickelt. Dieses Modell basiert auf der bereits angeführten Annahme, dass jede Person von Grund auf ein positives Bild von sich hat und danach bestrebt ist dieses Selbstbild zu erhalten. Schmader et al. (2008) erklären, dass Stereotype Threat durch ein *kognitives Ungleichgewicht* entsteht. Diese Unausgewogenheit tritt ein, wenn das eigentliche, positive Selbstbild der Person nicht mit den vorherrschenden, negativen Stereotypen, ihre soziale Gruppe betreffend, übereinstimmt. Bei Frauen entsteht dieses kognitive Ungleichgewicht durch die beiden konkurrierenden Identitäten, einerseits durch das positive Selbstbild und die Überzeugung über gute mathematische Fähigkeiten zu verfügen und andererseits durch die Konfrontation mit der Rolle der weiblichen, geschlechtsspezifischen Identität, die das negative Klischee von Frauen beinhaltet, über schwache mathematische Fertigkeiten zu verfügen. Dieses Ungleichgewicht spiegelt sich in der

gegensätzlichen und mehrdeutigen Information wider, die Frauen bezüglich ihrer geschlechtsspezifischen Identität, der Rolle als Frau und ihrem Selbstbild, in der jeweiligen Situation erhalten. Ohne Aktivierung der geschlechtsspezifischen Identität nehmen sich Frauen als fähig und kompetent wahr. Werden jedoch negative Stereotype über Frauen und deren schlechte mathematische Leistung durch die Aktivierung dieser weiblichen, geschlechtsbezogenen Identität zugänglich gemacht, zeigen sich klare Anzeichen von Stereotype Threat und eine daraus resultierende defizitäre Leistung. Rydell et al. (2009) demonstrierten im Zuge ihrer Untersuchung, dass die Vorgabe eines positiven und für das Selbstbild relevanten Stereotyps, wie zum Beispiel „CollegestudentInnen sind gut in Mathematik“, die Zugänglichkeit der weiblichen *College-Identität* erhöhte und gleichzeitig den Zugriff auf die geschlechtsspezifische Identität und Rolle der Frau, hemmte. Dadurch konnten sowohl die Beeinträchtigungen des Arbeitsgedächtnisses als auch die Defizite in der mathematischen Leistung eliminiert werden.

Die angeführten Studien von Schmader (2002) und Rydell & Beilock (2009) demonstrieren, dass der Effekt von Stereotype Threat durch den Grad der Identifikation mit dem eigenen Geschlecht, der Rolle als Frau und durch den Abruf der jeweiligen sozialen Identitäten, entweder verstärkt oder vermindert werden kann. Insbesondere der Abruf jener sozialen Identität der Frau mit dem integrierten positiven Selbstbild, konnte die Wirksamkeit von Stereotype Threat und daraus resultierende Verschlechterungen in den mathematischen Leistungen, unterbinden.

2.5.6 Stereotype Threat und *Stigmabewusstsein*

Ein weiterer, relevanter Aspekt in diesem Zusammenhang ist das Vorhandensein von individuellen Unterschieden der stereotypierten Personen in ihrem *Stigmabewusstsein*, worunter Pinel (1999), die individuell variierende und empfundene Selbstsicherheit in Bezug auf die zugehörige stigmatisierte Gruppe, versteht. Pinel (1999) entwickelte ein Instrument, den *Stigma consciousness questionnaire* (SCQ), mit dem diese individuellen Unterschiede im

Stigmabewusstsein gemessen werden können. Der SCQ wurde so konzipiert, dass er für den Einsatz für jede beliebige stereotypierte Gruppe modifiziert werden kann. Es wurden von Pinel selbst diesbezüglich Validierungsstudien durchgeführt, die zeigten, dass das Konstrukt des Stigmabewusstseins positiv mit dem öffentlichen und dem privaten Selbstbewusstsein (Fenigstein, Scheier & Buss, 1975) sowie mit der Geschlechtsrollenkonformität (Spence, Helmreich & Sawin, 1980) und negativ mit dem interpersonellen Vertrauen korreliert. Aufbauend auf seinen vorherigen Untersuchungen zum Stigmabewusstsein strebten Brown und Pinel (2003) mit ihrer Studie an, den Nachweis zu erbringen, dass Personen mit einem großen und erhöhten Stigmabewusstsein anfälliger und stärker auf die negativen Effekte des Stereotype Threat reagieren als Personen mit einem gering ausgeprägten Stigmabewusstsein. Im Rahmen ihrer Untersuchung konnten sie diesen Nachweis erfolgreich erbringen und zeigten, dass Frauen mit einem hohen Stigmabewusstsein unter der Stereotype Threat-Bedingung schlechtere mathematische Leistungen erbrachten als Frauen mit einem niedrigen Stigmabewusstsein unter derselben Bedingung. Darüber hinaus konnte demonstriert werden, dass Frauen mit hohem Stigmabewusstsein auch höhere Werte in der Geschlechtsrollenidentifikation (Schmader, 2002) aufwiesen. Die eindeutigen Effekte der Geschlechtsrollenidentifikation, die Schmader (2002) berichtete, konnten jedoch nicht klar repliziert werden. In Bezug auf potentielle Mediatorvariablen wie Motivation, Anstrengung und selbst wahrgenommene mathematische Fähigkeit, die im Zusammenhang mit Stereotype Threat stehen, konnten leider keine eindeutigen Ergebnisse erbracht werden. Zusammenfassend ist zu sagen, dass es in Bezug auf die Bedeutung und Gewichtung der Mediatorvariablen, die die Effekte des Stereotype Threat mit beeinflussen, sehr gegensätzliche Forschungsergebnisse gibt.

2.5.7 Stereotype Threat- negative Gedanken, Angst und Besorgnis

Cadinu, Rosabianca und Kiesner (2005) widmeten sich in ihrer Studie der Untersuchung des potentiellen Einflusses von negativen Gedanken auf die durch Stereotype Threat begünstigten und hervorgerufenen Leistungsdefizite. Es konnte

gezeigt werden, dass Frauen unter dem Druck von Stereotype Threat eine erhöhte Anzahl an negativen Gedanken in Bezug auf Mathematik-Tests aufwiesen, im Gegensatz zu ihren Kolleginnen, die nicht dem Stereotype Threat ausgesetzt wurden. Candinu et al., (2005) demonstrierten, dass auf die mathematische Leistung bezogene, negative Gedanken, einen potenziellen Verursacher für die verminderten Leistungen im Test unter Stereotype Threat darstellen.

Der Einfluss der Variablen Angst und Besorgnis auf die erbrachte Testleistung unter Stereotype Threat konnte in der Literatur nicht eindeutig nachgewiesen werden (Steele & Aronson, 1995; Aronson et al. (1999). In der bereits erwähnten Studie von Spencer et al. (1999) zeigten sich marginale Effekte der Angst und Besorgnis, die negative Auswirkungen auf die Leistung hatten. Tatsache ist, dass Ängstlichkeit und Besorgnis eine gewisse Rolle in Verbindung mit dem Phänomen Stereotype Threat spielen. Diesbezüglich ist die Forschung gefragt weitere Untersuchungen durchzuführen, die den Einfluss dieses Konstrukts operationalisieren.

2.5.8 Stereotype Threat – Auswirkungen auf das Erregungsniveau

Im Zusammenhang mit der Angst und Besorgnis ist auch der Aspekt des Erregungsniveaus zu erwähnen. Die von den betroffenen Personen wahrgenommene Bedrohung im Zuge der Stereotype Threat Aktivierung beeinflusst nachweislich deren Erregungsniveau. Ben-Zeev, Fein und Inzlicht (2004) beschäftigten sich im Rahmen ihrer Untersuchung mit dieser Thematik und stellten fest, dass Stereotype Threat eine gewisse Erregung hervorruft. Die verursachte, situationsspezifische Erregung und deren Attribuierung bzw. Ursachenzuschreibung haben eine negative Auswirkung auf die Testleistung. Es konnte somit nachgewiesen werden, dass durch das situationsspezifische Vorhandensein negativer Stereotype Erregung bei stereotypierten Personen hervorgerufen wird. Sowohl die Erregung an sich als auch die Art und Weise, wie diese Erregung von der Person interpretiert wird, wirken sich beeinträchtigend auf die Leistung aus.

2.5.9 Stereotype Threat in Verbindung mit der Erwartungshaltung

Einen weiteren Beitrag zur Erklärung von Stereotype Threat und dem zugrundeliegenden Prozess lieferten Cadinu, Maass, Frigerio, Impagliazzo und Lationtti (2002), indem sie sich mit der Rolle der Erwartung und der Erwartungshaltung einer Person in Bezug auf ihre mathematische Leistung auseinandersetzten. Es zeigte sich, dass Frauen, die das Fach der Mathematik als wichtig für sich selbst erachteten und die negative Manipulation erhielten „Frauen sind in Mathematik schlechter als Männer“, erbrachten nachweislich schlechtere Leistungen als ihre Kolleginnen in der Kontrollgruppe („Frauen und Männer sind gleich gut in Mathematik) und in der Gruppe, die die positive Information „Frauen sind in Mathematik besser als Männer“ erhielten. Frauen, die mit negativen Stereotypen konfrontiert wurden zeigten sowohl eine deutliche Verschlechterung ihrer mathematischen Leistung als auch eine verminderte Erwartungshaltung in Bezug auf das erfolgreiche Bearbeiten der Aufgaben. Die Ergebnisse dieser Untersuchung sprechen dafür, dass die verminderte Erwartungshaltung unter Stereotype Threat unter anderem eine Erklärung für die beeinträchtigte und defizitäre Leistung ist.

Eine Vielzahl an wissenschaftlichen Untersuchungen bezieht sich auf den Einflussfaktor des sozialen Geschlechts im Zusammenhang mit Stereotype Threat. In den bereits angeführten Studien wird hauptsächlich von der Bedeutung und dem Ausmaß der Identifikation mit dem sozialen Geschlecht gesprochen. Das Hauptaugenmerk liegt auf der sozialen, weiblichen Identität und der (gesellschaftlichen) Rolle der (als) Frau.

2.5.10 Stereotype Threat und die Rolle der Hormone

Halpern und Tan (2001) hingegen verfolgen in ihrer Studie einen biologischen Ansatzpunkt. Sie liefern eine weitere interessante Beleuchtung der Thematik, im Hinblick auf mögliche Zusammenhänge und Einflussfaktoren, indem sie sich mit der Rolle und Bedeutung der Hormone auseinandersetzen. Sie untersuchten den

Einfluss von Hormonen in der Menstruationsphase (Östradiol, Progesteron) auf kognitiven Fähigkeiten und Leistungen von türkischen StudentInnen.

Im Rahmen ihrer Untersuchung setzten Halpern und Tan (2001) ein selbstkonstruiertes Questionnaire, das *Beliefs About the Cognitive Abilities of Females and Males Questionnaire* ein, um Überzeugungen und Einschätzungen über die kognitiven Fähigkeiten von Männern und Frauen zu operationalisieren und daraus geschlechtsspezifische Stereotype abzuleiten.. Im Zuge dieses Questionnaires wurden die türkischen StudentInnen dazu aufgefordert Wahrscheinlichkeiten zu schätzen und anzugeben, ob sie glauben, dass die angeführte kognitive Fähigkeit eher Frauen oder Männern zuzuordnen ist. Die geschätzten Werte für beide Geschlechter sollten in Summe 100% ergeben. Darüber hinaus wurde den StudentInnen ein Selbstrating zur Selbsteinschätzung bestimmter kognitiver Fähigkeiten vorgegeben. Des Weiteren wurden drei kognitive Tests vorgegeben, der *Finding A's-Test* (Kimura & Hampson, 1994), repräsentativ für die weiblichen Stärken, der *Mental Rotation Test* (Vanderberg & Kuse, 1978) für die männlichen Stärken und der *Cattell Nonverbal Test of Intelligence* (Cattell & Cattell, 1963) als *neutraler* Test. Zusätzlich wurden Blutproben entnommen, um den Östradiolspiegel zu erfassen, in Anlehnung an die Annahme, dass ein gewisser Östradiolspiegel die Leistungen von Frauen bei Raumvorstellungsaufgaben verbessert. Die Resultate des Questionnaires konnten bei türkischen StudentInnen weder geschlechtsspezifische Stereotype noch Stereotype Threat Aktivierung ableiten, im Gegensatz zu amerikanischen StudentInnen in anderen Untersuchungen. Die Ergebnisse bezüglich der drei kognitiven Tests zeigten, dass ein erhöhter Östradiolspiegel zu geringen Verbesserungen in der Raumvorstellungskraft bei Frauen führte. Dies erwies sich konträr zu Erwartungen der Autoren, die davon ausgingen, dass ein erhöhter Östradiolspiegel einen mindernden Einfluss auf die räumlich-visuelle Raumvorstellung hat (Halpern & Tan, 2001). Im Zuge dieser Untersuchung konnte zwar festgestellt werden, dass Hormone sehr wohl eine relevante Rolle für die kognitive Leistung spielen, jedoch konnte nicht geklärt werden in welchem Ausmaß sie sich auf die Leistung auswirken. Das von Halpern und Tan (2001)

vorgestellte Questionnaire wurde für die vorliegende Studie adaptiert und im Rahmen der Testung vorgegeben

Zusammenfassend kann gesagt werden, dass Stereotype Threat durch verschiedene der angeführten Faktoren wie Erwartungshaltung, Bereichsidentifikation, Verfügbarkeit und Abruf sozialer Identitäten, ineffiziente Problemlösestrategien und negative Gedanken beeinflusst und moduliert wird und unter spezifischen Gegebenheiten aktiviert oder gehemmt werden kann. Darüber hinaus konnten Forschungsergebnisse belegen, dass die Aktivierung von Stereotype Threat eine vermindernde Wirkung auf die kognitiven Leistungen von Frauen im Speziellen hat.

Kapitel 3 Risikoverhalten

Im folgenden Kapitel wird genauer auf das Konstrukt des Risikoverhaltens und der Risikobereitschaft eingegangen. Zu Beginn findet eine kurze und allgemeine Auseinandersetzung mit dem Themenbereich der Entscheidungen statt, in der erläutert werden soll, was Entscheidungen grundsätzlich sind. Hauptaugenmerk liegt im Rahmen dieser Arbeit bei den Entscheidungen unter Risikobedingungen. Die Beschreibung von gängigen Entscheidungsmodellen und Theorien, die vorwiegend aus dem Bereich der Wirtschaftspsychologie stammen, sowie der Einbezug von geschichtlichen und theoretischen Hintergründen, sollen eine gewisse Grundlage für das Verständnis von Entscheidungen unter Risikobedingungen bieten. Des Weiteren soll das Konstrukt des Risikoverhaltens unter dem Aspekt geschlechtsspezifischer Unterschiede beleuchtet und anhand der Darstellung unterschiedlicher Forschungsergebnisse veranschaulicht werden. Abschließend wird der Game of Dice Task (GDT), ein computerisierter Test zur Erfassung der Risikofreude und Risikobereitschaft, vorgestellt. Es handelt sich dabei um ein objektives Verfahren, das auch im Zuge unserer Untersuchung eingesetzt wird.

Einleitend ist anzumerken, dass das Risikoverhalten insbesondere in der Wirtschaftspsychologie im Zuge der Erklärung und Erforschung von Kaufentscheidungen thematisiert wird und Gegenstand von wissenschaftlichen Untersuchungen ist. In der Literatur findet man verschiedene Studien, die belegen, dass Männer in gewissen Situationen risikofreudiger entscheiden als Frauen. Diese Ergebnisse und Überlegungen werden im Rahmen dieser Arbeit aufgegriffen und im Zusammenhang mit dem Antwort- und Entscheidungsverhalten von Frauen und Männern in einer Testsituation beleuchtet. Ein wichtiger Aspekt, der im Zusammenhang mit dem Risikoverhalten zu nennen ist, ist das Rateverhalten in einem Test.

Im vorherigen Kapitel wurde das Phänomen des Stereotype Threat in Bezug auf geschlechtsspezifische Unterschiede und unterschiedliche Einflussfaktoren näher erläutert. In diesem Kapitel liegt der Fokus auf dem Risikoverhalten, das neben dem Konstrukt Stereotype Threat einen weiteren Aspekt darstellt, der gewisse

Effekte auf die Testleistung hat und die Leistung bei der Bearbeitung spezifischer kognitiver und sozial-kognitiver Aufgaben modelliert und beeinflusst.

3.1. Historische und theoretische Hintergründe

Der Bereich der Entscheidungsforschung ist ein noch ziemlich junges Forschungsfeld, dessen Wurzeln ursprünglich in den Wissenschaftsbereichen der Philosophie, Ökonomie und Mathematik anzusiedeln sind. Die Anfänge der Entscheidungsforschung gehen zurück auf John von Neumann und Oskar Morgenstern, die 1947 in ihrem Buch „Spieltheorie und wirtschaftliche Verhalten“ erstmals eine Theorie zu präferentiellen Entscheidungen vorstellten (Jungermann et al., 2005)..

Der Bereich des Entscheidungsverhaltens ist eine weitgefächerte Thematik, die in verschiedenen Wissenschaftsdisziplinen behandelt wird und zu der es unterschiedliche Entscheidungstheorien gibt, die ihrerseits auf unterschiedlichen Menschenbildern basieren. Wie bereits eingangs erwähnt, setzt sich insbesondere der Bereich der Wirtschaftspsychologie mit Entscheidungen in unterschiedlichen Situationen wie zum Beispiel Gruppen- und Kaufentscheidungen sowie unter Ambiguitäts- oder Risikobedingungen, auseinander. Auf Basis deskriptiver und präskriptiver Entscheidungsmodelle wird das Entstehen einer Entscheidung von unterschiedlichen Ansatzpunkten erklärt und diskutiert. In Bezug auf das für diese Diplomarbeit relevante Konstrukt des Risikoverhaltens ist besonders die *Prospect Theory* von Kahneman und Tversky (1979) interessant, da sich diese Theorie explizit mit der Thematik der Risikoentscheidungen beschäftigt.

3.2. Was sind Entscheidungen?

Mit der Anforderung eine Entscheidung zu treffen, werden wir tagtäglich konfrontiert. Wir entscheiden uns, ob bewusst oder unbewusst, manchmal auf gründlichen Überlegungen basierend, teils aus Unwissenheit oder mangels Information wie bei Ambiguitätsentscheidungen (Kirchler, 2003) und manchmal auch im Bewusstsein über mögliche negative Konsequenzen, wie im Falle von

Risikoentscheidungen (Jungermann, Pfister & Fischer, 2005; Kirchler, 2003)
Eine Entscheidung zu treffen stellt somit eine wichtige Funktion in unserem Leben dar.

In der Entscheidungsforschung wird davon ausgegangen, dass sich eine Person in einer bestimmten Situation mindestens zwischen zwei Optionen für eine Alternative entscheiden muss. Diese Optionen können entweder Objekte (z.B. Auto, Wohnung) oder Handlungen (ins Kino gehen, einkaufen, Anweisungen an eine Person etc.) darstellen (Jungermann et al., 2005).

Laux (2003, S.1) definiert Entscheidung als,, die (mehr oder weniger bewusste) Auswahl einer von mehreren möglichen Handlungsalternativen“.

Eine weitere Definition von Jungermann, Pfister und Fischer (2005) versteht den Begriff der Entscheidung als einen Prozess, der durch zwei wesentliche Komponenten, Beurteilungen (judgements) und Wahlen (choices) charakterisiert wird. Sie beschreiben, dass der Ausgangspunkt dieses Prozesses die Diskrepanz zwischen dem tatsächlichen und dem gewünschten Zustand ist und die Person in weiterer Folge dazu veranlasst eine adäquate Option zu suchen, um diese Diskrepanz zu beseitigen. Der Prozess endet entweder mit der endgültigen Auswahl und Festlegung auf die jeweilige Alternative oder mit der Umsetzung und Ausführung der Option bzw. der retrospektiven Bewertung der Entscheidung (Jungermann et al., 2005).

Wie nun Entscheidungen zustande kommen, soll im folgenden Abschnitt anhand der Vorstellung unterschiedlicher Entscheidungsmodelle erläutert werden.

3.3. Wie entstehen Entscheidungen?

Innerhalb der Entscheidungsforschung werden zwei grundlegende, einander ergänzende Richtungen vertreten, die präskriptive und die deskriptive Entscheidungstheorie (Jungerman et al., 2005; Kirchler, 2003), die Erklärungen und Ansätze für das Zustandekommen von Entscheidung liefern.

3.3.1 Präskriptive Entscheidungstheorie

Bei der präskriptiven Entscheidungstheorie liegt der Fokus darauf, wie sich eine Person unter Berücksichtigung der Annahme des rationalen Denkens entscheiden sollte. Sie zielt darauf ab, Menschen bei schwierigen und komplexen Entscheidungen zu unterstützen. Dieser Ansatz bedient sich normativer Modelle, um Strategien und Methoden anzubieten, die die Person im Rahmen des Entscheidungsprozesses unterstützen sollen eine optimale Entscheidung zu treffen (Kirchler, 2003; 2008; Jungermann et al., 2005). Des Weiteren stellt sie formalisierte Regeln und Verfahren zur Strukturierung und Verarbeitung von Information zur Verfügung (Eisenführ & Weber, 1994). Präskriptive Entscheidungsmodelle basieren auf dem Menschenbild des „homo oeconomicus“, welches von der Annahme eines rational und vernünftig denkenden Menschen ausgeht, der sich aller Alternativen und deren Konsequenzen bewusst ist und danach die für ihn optimalste Option wählt (Kühberger, 1994).

3.3.2 Deskriptive Entscheidungsmodelle

Die deskriptive Entscheidungstheorie versucht im Gegensatz zu der präskriptiven Entscheidungstheorie, die von einem Optimum ausgeht, empirisch vorherherzusagen, wie sich eine Person in der Realität tatsächlich entscheidet (Kirchler, 2003). Im Rahmen der deskriptiven Entscheidungstheorie werden Theorien und Methoden angewendet, die das tatsächliche Verhalten in Entscheidungssituationen abbilden und empirisch überprüfen. Aufgrund der begrenzten zeitlichen Kapazität und der limitierten kognitiven Informationsverarbeitung wird von einem beschränkt rationalen Verhalten ausgegangen (Jungermann et al., 2005).

3.3.3 Prospekt Theory (PT)

Die Prospect Theory von Kahneman und Tversky (1979) illustriert den Einfluss subjektiver Einschätzungen auf das Verhalten des Entscheiders in Risikoentscheidungen, die durch die jeweilige Repräsentation des Problems

entweder auf einen Gewinn oder auf einen Verlust hingelenkt werden (Kirchler, 2003). Kahneman und Tversky (1979) erweiterten den klassischen entscheidungstheoretischen Ansatz und postulierten zwei Phasen im Entscheidungsprozess. Die erste Phase beinhaltet das Editieren eines vorhandenen Problems, worunter die Autoren verstehen, dass die situationsspezifischen Alternativen analysiert, enkodiert, transformiert und anschließend in einer vereinfachten mentalen Repräsentation dargestellt werden. In Abhängigkeit von der Beschreibung stellt sich die Option relativ zu der Ausgangslage des Entscheiders, entweder als Gewinn oder Verlust dar (Kirchler, 2003). In dieser Phase wird der Art und Weise der Problempäsentation, dem semantischen Rahmen, auch *framing* genannt, eine besondere Bedeutung beigemessen, da dieser die Aufmerksamkeit des Entscheiders auf einen Gewinn oder einen Verlust lenkt. Für die Entscheidung ist somit die Ausgangslage, der sogenannte *Referenzpunkt* ausschlaggebend. Darüber hinaus ist der Aspekt der Präferenzen des Entscheiders entscheidungsrelevant, denn diese sind je nachdem, ob eine Alternative einen Gewinn oder einen Verlust darstellt, unterschiedlich (Kirchler, 2003). In der zweiten Phase werden die editierten Alternativen evaluiert, was bedeutet, dass jede Alternative auf ihren subjektiven Wert bezogen bewertet und gemäß der subjektiven Aussichten auf Gewinn oder Verlust, eine Alternative ausgewählt wird (Kirchler, 2003; Jungermann et al., 2005).

Im Rahmen der PT wird davon ausgegangen, dass die Entscheidung einerseits von den Konsequenzen einer Alternative und andererseits von deren Wahrscheinlichkeiten determiniert wird. Die Konsequenzen einer Alternative werden von dem Entscheider subjektiv transformiert, verändert repräsentiert und anschließend relativ zu einem spezifischen Referenzpunkt bewertet. Dieser Referenzpunkt kann unabhängig davon, ob eine Entscheidung zu gleichen Konsequenzen führt, unterschiedlich sein (Jungermann et al., 1999).

Ein Beispiel zur Verdeutlichung: Ein Läufer bereitet sich für einen Triathlon vor. In der Laufdisziplin erwartet er sich sehr gute Ergebnisse und geht davon aus unter den schnellsten 5 zu sein. Eine Platzierung unterhalb der besten 5 würde ihn enttäuschen. In der Schwimmdisziplin hingegen wäre er über eine Platzierung

unterhalb der besten 5 überglücklich. Dieses Beispiel zeigt, dass der Referenzpunkt relevant für die Beurteilung einer Konsequenz ist und dass objektiv gleiche Konsequenzen, je nach Referenzpunkt, einen unterschiedlichen subjektiven Wert für eine Person haben können. Im Zusammenhang mit dem Referenzpunkt ist auch der *Framing Effekt* (Kirchler, 2003, Jungerman et al.2005) anzuführen, der erklärt, dass je nachdem, ob die Problempräsentation die Aufmerksamkeit auf negative Konsequenzen oder positive Konsequenzen richtet, sich der Referenzpunkt verschiebt und verändert. Verschlechterungen relativ zum Referenzpunkt werden als Verluste verbucht und im Vergleich zu Gewinnen besonders intensiv und negativ wahrgenommen. Es zeigt sich, dass Menschen dann die riskante Entscheidung gegenüber der sicheren bevorzugen, wenn die Optionen und Konsequenzen positiv dargestellt (*geframt*) werden. Ein Beispiel: Ein neues Medikament gegen Bluthochdruck wird auf den Markt gebracht, das über schwere Nebenwirkungen verfügt. Es wird aber als das beste und wirkungsvollste Blutdruckmedikament propagiert, das bereits vielen Betroffenen geholfen hat. In diesem Fall wird sich die betroffene Person risikoreich entscheiden, da der Fokus auf den positiven Konsequenzen liegt. Wird derselbe Fall anders dargestellt und der Fokus auf die schwerwiegenden Nebenwirkungen gelegt, die die Lebensqualität massiv einschränken, wird die betroffene Person die sichere Entscheidung gegenüber der riskanten Entscheidung bevorzugen.

Kahnemann und Tversky (1979) stellen diesen Vorgang mittels einer sogenannten *Wertfunktion* dar. Demnach werden Konsequenzen, die oberhalb des Referenzpunktes liegen, als Gewinne angesehen, dessen Wertfunktion in diesem Fall konkav ist. Werden Konsequenzen unterhalb des Referenzpunktes angesiedelt, dann werden sie als Verluste verbucht und die dazugehörige Wertfunktion verläuft konvex. Mittels konkaver und konvexer Wertfunktion kann nun eine Aussage über das Risikoverhalten des Entscheiders gemacht werden (Jungermann et al., 2005) Die konkave Funktion soll die Risikoscheue eines Entscheiders bei Gewinnchancen veranschaulichen. Es zeigt sich, dass der subjektive Wert bei geringen Beträgen schneller wächst als die Höhe des eigentlichen Gewinns. Anhand des Lotterie-Beispiels lässt sich darstellen, dass eine Person bei der Wahl zwischen einem sicheren Gewinn von 70 Euro und der

Möglichkeit bei der Lotterie mit einer Wahrscheinlichkeit von 0,85 100 Euro und mit einer Wahrscheinlichkeit von 0,15 gar nichts zu gewinnen, sich für den sicheren Gewinn entscheidet. Der subjektive Wert der Lotterie wird geringer als der subjektive Wert des sicheren Gewinns von 70Euro wahrgenommen. Im Falle der konvexen Funktion demonstriert die Risikofreude- bzw. suche bei drohenden Verlusten. Im Lotteriebeispiel lässt sich die Situation wie folgt darstellen: Eine Person kann entweder einen sicheren Verlust von 70 Euro erleiden oder in der Lotterie mit einer Wahrscheinlichkeit von 0,85 100 Euro verlieren oder mit einer Wahrscheinlichkeit von 0,15 nichts verlieren. Die Person entscheidet sich für die Lotterie, da der subjektive Wert der Lotterie als größer eingestuft wird als der des sicheren Verlustes. Der Unterschied, im Falle des Verlustes in der Lotterie, wird als nicht so groß empfunden wie der sichere Verlust, da bei der Lotterie noch immer die Möglichkeit besteht mit einer Wahrscheinlichkeit von 0,15 nichts zu verlieren.

Im Rahmen der PT ist noch kurz eine zweite Funktion, die *Entscheidungsgewichtungsfunktion* zu nennen, die Auskunft darüber gibt, welches Gewicht der Entscheider einer bestimmten Wahrscheinlichkeit bei einer Entscheidung zuteilt. Es zeigt sich, dass Personen kleinen Wahrscheinlichkeiten ein überproportional großes Gewicht beimessen (Jungermann et al., 2005). Ein Beispiel: Menschen haben größere Angst mit einem Flugzeug abzustürzen als mit einem Auto zu verunglücken, obwohl das Flugzeug im Vergleich zu den anderen Verkehrsmitteln noch als „Sicherste“ angesehen werden kann. Die Person weist der geringen Wahrscheinlichkeit eines Flugzeugabsturzes ein überproportional großes Gewicht zu.

Zusammenfassend kann man sagen, dass die PT einen differenzierten Erklärungsansatz und Einblick in das Entscheidungsverhalten einer Person liefert, in dem sie unterschiedliche Umgebungs-, Einfluss- und Störfaktoren, die Auswirkungen auf den Entscheidungsprozess, auf das gesamte Entscheidungsverhalten und letztendlich auf die endgültige Entscheidung des Entscheiders haben, berücksichtigt.

Anschließend an diese Darstellung folgt eine Erläuterung und Beschreibung von Entscheidungen unter Risikobedingungen.

3.4. Entscheidungen unter Risikobedingungen

Entscheidungen unter Risikobedingungen zeichnen sich dadurch aus, dass die Person über die nötige Information bezüglich der Auftrittswahrscheinlichkeit negativer Konsequenzen verfügt.

Bevor jedoch noch etwas differenzierter auf diese Thematik eingegangen wird, wird der Begriff des Risikos kurz erläutert. In der Literatur gibt es für den Begriff Risiko mehrere unterschiedliche Definitionen und Ansätze. Es zeigt sich jedoch, dass es eine grundlegende und sehr allgemeine Definition gibt, die alle wissenschaftlichen Disziplinen gemeinsam haben. Unter dem Begriff des Risikos wird die Beschreibung eines Ereignisses verstanden, das mit dem möglichen Eintreten von negativen Konsequenzen verbunden ist (Kirchler, 2003; Jungerman et al., 2005, Breakwell, 2007). Risiko wird im Allgemeinen anhand zweier Dimensionen beschrieben und definiert. Die erste Dimension bezieht sich auf Wahrscheinlichkeiten und die zweite Dimension betrifft Konsequenzen (Breakwell, 2007).

Breakwell (2007) definiert Risiko demnach als die Wahrscheinlichkeit des Eintretens eines spezifischen negativen Ereignisses zu einem bestimmten Zeitpunkt. Das Risiko ist somit die Wahrscheinlichkeit eines bestimmten und unerwünschten Ereignisses, die aus der Konfrontation mit einer Gefahr, einem Wagnis resultiert (Breakwell, 2007). Entscheidungen unter Risiko werden in der Literatur auch als Entscheidungen unter Unsicherheit bezeichnet, da bei diesen Entscheidungen die Konsequenzen unsicher sind (Jungermann et al., 2005). Im Rahmen des Entscheidungsprozesses unter Risikobedingungen kann sich eine Person entweder risiko-aversiv beziehungsweise risiko-vermeidend oder risiko-freudig beziehungsweise risiko-suchend verhalten.

Eine Person entscheidet sich risikoreich, wenn sie sich trotz des Wissens über die möglichen unerwünschten und unangenehmen Konsequenzen für die jeweilige

Option entscheidet (Jungerman et al., 2005; Kirchler, 2003). Ein Beispiel dafür: sich dafür zu entscheiden zu einer Prüfung zu gehen, obwohl man weiß, dass man kaum gelernt hat. Die Wahrscheinlichkeit die Prüfung nicht zu bestehen ist groß, aber es besteht trotzdem eine geringe Wahrscheinlichkeit die Prüfung doch zu bestehen. Die negative Konsequenz wäre das Nicht-Bestehen der Prüfung und die positive Konsequenz wäre das Bestehen mit einer schlechten Note. Wenn die Konsequenz des Bestehens eine größere Wichtigkeit für die Person hat, dann entscheidet sie sich risikoreich, da sie nicht weiß, welche Konsequenz letztendlich eintritt, entweder das Bestehen oder Nicht-Bestehen der Prüfung. Bei Risikoentscheidungen handelt es sich somit um Entscheidungen mit einem ungewissen Ausgang (Kirchler, 1999). Im Zuge wirtschaftspsychologischer Untersuchungen konnte festgestellt werden, dass Menschen prinzipiell dazu neigen risiko-scheu zu agieren und den sicheren Gewinn präferieren. Risikofreudige Entscheidungen werden meist dann getroffen, wenn der subjektive Wert des erwarteten Gewinns jenen des Risikos übertrifft. Generell kann man sagen, dass Menschen sichere Entscheidungen den Risikoentscheidungen vorziehen und riskante Entscheidungen gegenüber Entscheidungen unter Ambiguität präferieren (Kirchler, 1999).

Zusammenfassend ist zu sagen, dass es sich bei Risikoentscheidungen um einen speziellen Entscheidungstypus handelt, dessen Auftreten von zwei wesentlichen Komponenten abhängt:

1. Mindestens eine der zur Wahl stehenden Optionen hat negative Konsequenzen.
2. Das Eintreten dieser negativen Konsequenzen ist unsicher.

Des Weiteren charakterisiert Huber (2004) Entscheidungen unter Risikobedingungen durch zwei weitere grundlegende Aspekte: 1. durch den subjektiven Wert der Konsequenzen und 2. durch die subjektive Wahrscheinlichkeit dieser Konsequenzen.

Zur Veranschaulichung bleiben wir bei dem Prüfungsbeispiel von vorher: Eine Person hat kaum gelernt und entscheidet sich trotzdem zur Prüfung zu gehen. Es kann mindestens eine Alternative gewählt werden, die negative Konsequenzen zur

Folge hat. 1. Konsequenz: die Person besteht die Prüfung nicht. 2. Konsequenz: Bestehen der Prüfung mit einer sehr schlechten Note. 3. Konsequenz: Bestehen der Prüfung mit einer guten Note. Welche der Konsequenzen in dieser Situation eintritt ist unsicher. Bei der Entscheidung unter Risikobedingungen spielen der subjektive Wert dieser Konsequenzen und deren subjektive Wahrscheinlichkeiten eine bedeutende Rolle. Es hängt davon ab, welchen subjektiven Wert die Person den Konsequenzen zuteilt. Es stellt sich somit die Frage wie wichtig das Bestehen der Prüfung für die Person ist und welche Chancen sie sich ausrechnet, dieses Ziel auch wirklich zu erreichen. Ist der Person das alleinige Bestehen der Prüfung wichtig, egal mit welcher Note, wird sie der Konsequenz des Bestehens mit einer schlechten Note den größten subjektiven Wert zuteilen und dementsprechend auch die subjektive Wahrscheinlichkeit des Eintretens dieser Konsequenz einschätzen.

Risikoentscheidungen stellen einen komplexen Bereich dar und demonstrieren, dass Menschen sich nicht immer nach logischen und rationalen Regeln entscheiden. Menschen unterscheiden sich in ihrem Risikoverhalten dahingehend, dass sie risiko-scheu oder risiko-suchend eine Entscheidung treffen. Des Weiteren gibt es auch markante Geschlechtsunterschiede im Risikoverhalten auf die im folgenden Abschnitt eingegangen werden soll.

3.5. Geschlechtsspezifische Unterschiede im Risikoverhalten

In diesem Abschnitt soll die Thematik der geschlechtsspezifischen Unterschiede bei Frauen und Männern in Bezug auf das Risikoverhalten beleuchtet werden. In der Literatur gibt es mehrere Befunde dafür, dass Männer risikoreicher agieren als Frauen (Ben-Shakar & Sinai, 1991). Diesbezüglich zeigen mehrere Untersuchungsergebnisse, dass es einen klaren Zusammenhang zwischen dem Geschlecht und der Testleistung bei verschiedenen Multiple choice Tests gibt (Ben-Shakar & Sinai, 1991). Es zeigt sich, dass es geschlechtsspezifische Unterschiede im Antwortverhalten, insbesondere auf das Raten und risikoreiche Antworten bezogen, gibt.

Bevor eine genauere Betrachtung relevanter Forschungsergebnisse erfolgt, sollen einleitend allgemeine und bereits mehrmals untersuchte geschlechtsspezifische Unterschiede bezüglich kognitiver Fähigkeiten vorgestellt werden, da sich diese auch im Antwortverhalten abbilden können.

Wie bereits mehrfach erwähnt, belegt eine Vielzahl an Studien, dass es geschlechtsspezifische Unterschiede bei Männern und Frauen in der Bearbeitung bestimmter kognitiver Aufgaben gibt. Maccoby und Jacklin (1974) identifizierten geschlechtsspezifische Unterschiede speziell in drei Bereichen: a.) verbale Fähigkeit, b.) räumlich-visuelle Fähigkeit, c.) mathematische Fähigkeit. Männer zeichnen sich durch besser ausgeprägte räumlich-visuelle und mathematische Fähigkeiten aus (Maccoby & Jacklin, 1974, zitiert nach Ben-Shakhar & Sinai, 1991, S.23). Sie erzielen bessere Ergebnisse sowohl im Mental Rotation Test, im Embedded Figures Test als auch beim räumlichen Navigieren und Kartenlesen. Im Kindesalter spielen Knaben bevorzugt mit mechanisch-technischem Spielzeug und im Erwachsenenalter zeichnet sich eine Überlegenheit beim Lösen und Bewältigen technischer und physikalischer Probleme ab (Baron-Cohen et al., 2005). Die Stärke der Frauen hingegen liegt bei der verbalen Fähigkeit (Maccoby & Jacklin, 1974). Frauen erbringen bessere Leistungen in Tests, die Emotionserkennung, soziale Kompetenz, soziale Sensibilität und Wortflüssigkeit erfassen. Des Weiteren beginnen Mädchen früher zu sprechen als Burschen (Baron-Cohen et al., 2005). Diese geschlechtsspezifischen Unterschiede in den kognitiven Fähigkeiten und die damit einhergehenden Leistungsunterschiede bilden sich unter anderem auch im Antwortverhalten bei der Bearbeitung unterschiedlicher kognitiver Aufgaben ab.

Ben-Shakhar und Sinai (1991) bezogen sich in ihrer Untersuchung auf geschlechtsspezifische Unterschiede in der Tendenz, Items in einem Multiple Choice Test auszulassen und Antworten zu raten. Ziel der Studie war es, eine Verbindung zwischen den Unterschieden in der Testleistung und den Unterschieden im Antwortverhalten herzustellen. Des Weiteren vermuteten die Autoren, dass ein Teil der Leistungsunterschiede durch den Antwortstil der Person erklärbar sei und es eine klare Beziehung zwischen Geschlecht und

Ratetendenzen gibt. Diesbezüglich ist besonders sowohl das risikoreiche, risikofreudige Antwortverhalten als auch das Rateverhalten von Interesse und Bedeutung. Es wird angenommen, dass Männer risikofreudiger antworten als Frauen und somit einerseits auch eine höhere Tendenz zum Raten und andererseits geringere Anzeichen für Auslassungen aufweisen. In Bezug auf das Rateverhalten in Multiple Choice Tests ist zu erwähnen, dass es unterschiedliche Formen und Arten des Ratens gibt. Global wird zwischen zufälligem Raten und Raten, das auf einem Teilwissen basiert, unterschieden. Ratetendenzen werden durch unterschiedliche Faktoren beeinflusst, begünstigt und ausgelöst. Sie sind abhängig von situationsspezifischen Faktoren, der Testinstruktion, dem Zeitdruck, dem Inhalt und der Schwierigkeit der einzelnen Antworten/Items. (Thorndike, 1971, zitiert nach Ben-Shakhar & Sinai, 1991).

Die Ergebnisse der Studie bestätigen, dass Frauen dazu neigen mehr Items im Rahmen der Aufgabenbearbeitung auszulassen als Männer. Es ist interessant anzumerken, dass Frauen generell in mehreren Subtests eine klare Auslassungstendenz zeigten, unabhängig davon, ob in diesen Subtests geschlechtsspezifische Unterschiede beobachtet wurden. Der Unterschied im Rateverhalten kann somit diesbezüglich nicht nur alleine auf die bessere Testleistung der Männer in den meisten Subtests zurückgeführt werden. Des Weiteren zeigte sich, dass die zum Raten motivierende Instruktion keinen minimierenden und eliminierenden Effekt auf die geschlechtsspezifischen Unterschiede in der Ratetendenz hatte. Daraus lässt sich schließen, dass derartige Instruktionen, die zum Raten auffordern, nicht adäquat und unzureichend für die Eliminierung der Unterschiede im Rateverhalten sind (Ben-Shakhar & Sinai, 1991).

In ihrer Metaanalyse analysierten Byrnes, Miller und Schafer (1999) 150 Studien von insgesamt 500 gesichteten Studien, in denen Risikotendenzen und -entscheidungen von männlichen und weiblichen Versuchsteilnehmern untersucht und miteinander verglichen wurden. Die beiden Hauptziele dieser Untersuchung waren einerseits die Bestimmung einer durchschnittlichen Effektgröße von geschlechtsspezifischen Unterschieden im Risikoverhalten und andererseits

aufzuzeigen, in welchem Ausmaß diese geschlechtsspezifischen Unterschiede bezüglich Kontext und Alter variieren. Die Autoren kodierten alle 150 Studien nach dem Typ der Aufgabenstellung und Anwendung (hypothetische Wahl/Entscheidung, Selbsteinschätzung des Verhaltens und beobachtetes Verhalten), nach dem Aufgabeninhalt und nach dem Alter. Die Ergebnisse sind konsistent mit vergangenen Befunden und unterstützen die Annahme, dass Männer mehr dazu geneigt sind, sich risikoreicher bzw. -freudiger zu verhalten und zu entscheiden als Frauen. Die Mehrheit der Effekte (60%) in dieser Analyse spricht für eben diese Aussage, dass das Risikoverhalten bei Männern ausgeprägter ist als bei Frauen. Fast die Hälfte dieser Effekte (48%) war größer als .20 (.20 beschreibt einen kleinen Effekt, .50 einen mittleren und .80 einen starken Effekt, Bortz & Döring, 2006). Darüber hinaus stellten Byrnes et al. (1999) fest, dass Männer eine fehlende Einsicht und ein beeinträchtigtes Urteilsvermögen in Hinblick auf risikoreiche Entscheidungen aufwiesen. Es zeigte sich, dass Männer prinzipiell weitaus mehr risikoreiche Entscheidungen treffen und das auch in Situationen, in denen sich offensichtlich eine risikoreiche Entscheidung ganz und gar nicht bewährt. Im Gegensatz dazu konnte bei Frauen demonstriert werden, dass sie sogar in harmlosen Situationen oder in Situationen, in denen eine risikoreiche Entscheidung vorteilhaft wäre, eine Abneigung gegenüber risikoreichen Entscheidungen aufwiesen und sich dementsprechend risiko-vermeidend verhielten (Byrnes et al., 1999). Die Resultate dieser Metaanalyse sprechen für die Annahme, dass geschlechtsspezifische Unterschiede in Bezug auf Kontext und Alter variieren. Es zeigte sich, dass bestimmte Themen (z.B. Rauchen) bezogen auf das Alter gleichbleibende geschlechtsspezifische Unterschiede produzieren und andere Themenbereiche (z.B. Autofahren) hingegen eine entwicklungsbezogene Vergrößerung oder Verminderung der geschlechtsspezifischen Unterschiede verursachen. (Byrnes et al., 1999). Das Alter und der Entwicklungsaspekt sind somit interessante und relevante Faktoren, die im Zusammenhang mit geschlechtsspezifischen Unterschieden im Risikoverhalten zu berücksichtigen sind.

Zusammenfassend ist anzumerken, dass Individuen sich prinzipiell in ihren Verhaltens- und Denkweisen unterscheiden. Dies gilt, wie die berichteten

Ergebnisse zeigen, auch für den spezifischen Bereich des Risikoverhaltens. Personen zeigen in bestimmten Situationen verschiedene Verhaltensmuster und Herangehensweisen und treffen aufgrund unterschiedlicher Faktoren, Hinweisreizen, Umwelt und- Umgebungsbedingungen sowie aufgrund sozialer und individueller Gegebenheiten ihre Entscheidungen. Personen unterscheiden sich nicht nur allgemein in ihrem Risikoverhalten, sondern weisen auch Unterschiede in der Ausprägung ihrer Risikofreude und Risikobereitschaft auf.

3.6. Erfassung des Risikoverhaltens

Die Frage, die sich nun stellt ist: Wie wird Risikoverhalten und speziell die Risikobereitschaft erhoben und gemessen?

Im Folgenden wird ein kurzer Überblick über einige in der Literatur diskutierten Erhebungsinstrumente zur Erfassung der Risikobereitschaft gegeben.

In dem sehr aktuellen Artikel von Rubio, Hernández, Zaldívar, Márquez und Santacreu (2010) werden zwei Messinstrumente, der Roulette Test (RT) und der Crossing the Street Test (CtST), zur Erfassung und Operationalisierung der Risikobereitschaft vorgestellt und im Zuge ihrer Untersuchung validiert. In der Literatur werden einige Verfahren angeführt, die versuchen, individuelle Unterschiede im Risikoverhalten und in der Risikobereitschaft bzw. Risikofreude zu erfassen und abzubilden. Viele dieser Verfahren basieren jedoch auf der Selbsteinschätzung der Versuchsteilnehmer und sind damit verfälschbar und keinesfalls objektiv. Rubio et al. (2010) führen im Rahmen ihrer Überlegungen drei objektive Verfahren an, die den Verhaltensaspekt und den Einfluss von Persönlichkeitseigenschaften und – faktoren berücksichtigen und untersuchen.

- **Iowa Gambling Task** (IGT, Bechara, Damasio, Damasio & Anderson, 1994, zitiert nach Rubio et al., 2010): Ist ein computerisiertes Verfahren zur Erfassung von Entscheidungen unter Ambiguitätsbedingungen und Beeinträchtigungen bei der Entscheidungsfindung. Hiermit wird jedoch nicht die reine Risikobereitschaft erfasst. Es wird versucht mittels IGT reale Entscheidungen zu simulieren. Bei der IGT handelt es sich um ein

virtuelles Kartenspiel bei dem die Person eine Karte aus vier Kartendecks auswählen soll. Der Person wird gesagt, dass sie bei jeder Karte, die sie auswählt Geld gewinnt. Einige Male gewinnt die Person, doch nach einiger Zeit verliert die Person auch, wenn sie eine Karte auswählt. Es gibt somit „gute, gewinnbringende Kartendecks“ und „schlechte, verlustbringende Kartendecks“. Ziel ist es so viel Geld wie möglich zu gewinnen.

- **Ballon Analog Risk Task** (BART, Lejuez et al., 2002, zitiert nach Rubio et al., 2010): Mit diesem Instrument wird die reine Risikobereitschaft gemessen. Die Person sieht einen virtuellen Luftballon auf dem Bildschirm, den sie bis zum Maximum aufblasen, aber nicht zum Platzen bringen soll. Mit jedem Mal pumpen gewinnt die Person einen bestimmten Geldbetrag. Für die Ballone werden unterschiedliche Explosionswahrscheinlichkeiten vom ersten bis zum 128 zigsten Pumpversuch angegeben. Umso größer der Ballon aufgepumpt wird, desto größer ist der erzielte Gewinn.
- **Betting Dice Test** (BDT, Arend, Botella, Contreras, Hernández & Santacreu, 2003, zitiert nach Rubio et al., 2010): Dieses Verfahren geht davon aus, dass Risikobereitschaft ein stabiles Verhalten ist. Beim BDT wird die Person angewiesen auf eine von vier Optionen zu setzen, die aus dem Wurf von zwei Würfeln resultieren könnten. Je größer die Auftrittswahrscheinlichkeit für eine Option ist, desto kleiner ist der Gewinn. Die Person erhält kein Feedback nach den gesetzten Wetten, um somit die Beeinflussung durch die vorher gewählten Alternativen auszuschalten.

Es zeigt sich, dass objektive, verhaltensbasierte Messinstrumente geeignet dazu sind die Risikofreude und-bereitschaft zu erfassen. Die in der Studie untersuchten Verfahren RT und CtST sind valide Verfahren zur Messung der Risikobereitschaft und zur Vorhersage des Verhaltens von Personen in unterschiedlichen Lebenssituationen.

Abschließend ist zu sagen, dass es unterschiedliche objektive Verfahren gibt, die versuchen das Risikoverhalten adäquat abzubilden. Im Folgenden soll nun ein weiteres und sehr aktuelles objektives Verfahren zur Erfassung des Risikoverhaltens und im Speziellen der Risikobereitschaft vorgestellt werden, dass auch im Rahmen der vorliegenden Untersuchung zur Erfassung des Konstrukts des Risikoverhaltens bzw- der Risikobereitschaft- und freude eingesetzt wurde.

3.6.1 Game of Dice Task (GDT)

In diesem Abschnitt soll, wie bereits erwähnt, die Game of Dice Task (GDT) als Instrument zur Erfassung des Risikoverhaltens vorgestellt werden.

Der *Game of Dice Task* (GDT, Brand, Fujiwara et al. 2005) ist ein objektives Verfahren zur Erfassung des Risikoverhaltens und der Risikofreude. Es wurde entwickelt um Entscheidungen unter Risikobedingungen abzubilden.

Der *Game of Dice Task* (GDT) ist ein computerisiertes Würfelspiel, bei dem die Versuchsteilnehmer vor die Aufgabe gestellt werden, mit fiktivem Geld auf das Ergebnis eines Würfelwurfs wetten zu müssen, indem sie eine von mehreren Alternativen wählen, die sich hinsichtlich der Gewinnwahrscheinlichkeit und der Gewinn- bzw. Verlusthöhe unterscheiden. Die Versuchsteilnehmer starten mit einem Guthaben von 1000€ und müssen sich vor jedem der 18 Durchgänge entscheiden, ob sie auf eine, zwei, drei oder vier Augenzahlen setzen. Je nachdem, ob die Person nun auf eine, zwei, drei oder vier Augenzahlen setzt, erhöht oder reduziert sich ihre Gewinnchance und die Höhe des fiktiven Gewinns. Die Gewinnwahrscheinlichkeit und –höhe sind dabei reziprok miteinander verknüpft, so dass eine Diskrepanz zwischen hohen, aber unwahrscheinlichen und geringen, aber wahrscheinlichen Gewinnen entsteht. Um im GDT zu gewinnen, muss die Versuchsperson eine vorteilhafte Entscheidungsstrategie entwickeln. Für diese Anforderung und Fähigkeit wiederum ist der Bereich der exekutiven Funktionen essentiell. In einigen Studien (Brand et al., 2004; Brand et al., 2005;) konnte gezeigt werden, dass es einen Zusammenhang zwischen der Anzahl an risikoreichen Entscheidungen im GDT und den exekutiven Funktionen gibt. Für

die vorteilhafte und effiziente Bearbeitung des GDT scheinen insbesondere exekutive Funktionen wie die Planung, die Kategorisierung von Optionen, die Entscheidung von Prioritäten, die kognitive Flexibilität und die Verarbeitung von Rückmeldungen vorheriger Entscheidungen von Bedeutung zu sein.

In ihrer Studie untersuchten Euteneuer et al. (2009) Beeinträchtigungen und Dysfunktionen im Entscheidungsverhalten betreffend Ambiguitäts- und Risikoentscheidungen bei Menschen mit Parkinson. Im Rahmen dieser Untersuchung wurden die objektiven Verfahren, der Iowa Gambling Test (IGT) zur Erhebung von Entscheidungen in mehrdeutigen, ambivalenten Situationen mit impliziten Regeln und den Gambling Dice Task (GDT) zur Erfassung von Entscheidungen unter Risikobedingungen, angewendet. Die Autoren geben an, dass ihre Studie die erste wäre, in der das Entscheidungsverhalten von Personen mit Parkinsonerkrankung mittels beider, sowohl mit dem IGT als auch mit dem GDT untersucht worden ist. Es bestand die Annahme, dass Personen mit Parkinson größere Defizite und Beeinträchtigungen bei Entscheidungen unter Risikobedingungen aufweisen würden. Die Ergebnisse der Untersuchung bestätigen diese Annahme, da Personen mit Parkinson eine nachweislich schlechtere Leistung im GDT erbrachten als die Kontrollgruppe. Diese Resultate sind auch konsistent mit erbrachten Belegen in der Literatur, dass Personen mit Parkinson, Beeinträchtigungen im Bereich der exekutiven Funktionen aufweisen, die mitunter bedeutend für das Entscheidungsverhalten in risikoreichen Situationen sind. Es konnte erneut gezeigt werden, dass der GDT ein aussagekräftiges und objektives Instrument zur Erforschung und Erfassung der Risikofreude und -bereitschaft ist.

Zusammenfassend kann man sagen, dass der GDT ein adäquates und relevantes Verfahren zur Abbildung des Risikoverhaltens darstellt und deshalb im Zuge der vorliegenden Testung eingesetzt wurde.

Kapitel 4 Theory of Mind:

In diesem Kapitel soll der Begriff der *Theory of Mind (ToM)* geklärt und näher beleuchtet werden. Es folgt eine kurze und prägnante Auseinandersetzung mit dem geschichtlichen Hintergrund und den Ursprüngen der *Theory of Mind* sowie eine Aufarbeitung der wichtigsten Erklärungsansätze und Theorien statt. Abschließend werden interessante und relevante Forschungsergebnisse vorgestellt, um einen Einblick in den bisherigen Stand des Wissens zu geben.

4.1. Geschichte und Hintergrund

Die Ursprünge der entwicklungspsychologischen Erforschung der ToM gehen zurück auf die wissenschaftlichen Arbeiten und Tierexperimente von Premack und Woodruff (1978). Sie stellten im Rahmen einer ihrer Untersuchungen die provokante Frage: „Does the chimpanzee have a theory of mind?“ und gingen von der Annahme aus, dass theoretische Konzepte die Voraussetzung für die Fähigkeit, sich selbst und anderen mentale Zustände zuzuschreiben sind. Mentale Zustände sind demnach nicht beobachtbar und müssen somit erst erschlossen werden. Premack und Woodruff (1978) führten im Rahmen ihrer Untersuchung Experimente mit einer Schimpansin durch, der mittels Videosequenzen unterschiedliche Problemsituationen (z.B. eine Person, die versucht aus einem geschlossenen Käfig herauszukommen) gezeigt wurden, die sie durch das Auswählen der korrekten Lösung (z.B. das richtige Foto mit der Lösung auszuwählen) aus einer Anzahl an Alternativen meistern sollte. Die Schimpansin wählte nachweislich häufiger das richtige Foto aus, woraufhin Premack und Woodruff (1978) schlossen, dass Schimpansen fähig zur Zuschreibung mentaler Zustände seien. Diese Erkenntnis kann jedoch nicht ganz kritiklos übernommen werden, da die an die Schimpansin gestellten Aufgaben auch ohne die Zuschreibung mentaler Zustände bewältigbar waren und man infolgedessen nicht automatisch davon ausgehen darf, dass die Schimpansin die Fähigkeit der Zuschreibung mentaler Zustände für die Lösung herangezogen hat. Tatsächlich hätte die Schimpansin allein durch die Repräsentation der Problemsituation zu der

korrekten Lösung kommen können, ohne sich je dabei die Frage gestellt zu haben, wie ein anderes Individuum die Problemsituation repräsentiert. (Sodian, 2007)

Um zu prüfen, ob eine Person bereits die Fähigkeit der Zuschreibung mentaler Zustände besitzt, ist die Repräsentation einer falschen Überzeugung einer Person über eine Situation erforderlich. Die Vorhersage der Handlungen einer Person, die auf ihren wahren Überzeugungen basiert, ist diesbezüglich weder hilf- noch aufschlussreich, da dies auch ohne der Zuschreibung mentaler Zustände, nämlich allein aufgrund der Repräsentation der Welt möglich ist. (Sodian, 2007).

4.2. Definition und Beschreibung

Was ist die *Theory of Mind*? Der Begriff der *Theory of Mind* (ToM) ist ein mehrdeutiger Begriff für den mehrere und unterschiedliche Definitionsversuche existieren. Prinzipiell ist zu sagen, dass jeder Mensch im Laufe seines Lebens, in den frühen Kindheitsjahren, die *Theory of Mind*, ein gewisses Grundverständnis von sich selbst und unseren Mitmenschen, ausbildet.

Erstmals zum Einsatz kam die Bezeichnung der *Theory of Mind* bei Fodor (1978) und bei Premack und Woodruff (1978). Sie bezeichneten damit eine „spezielle geistige Leistung, nämlich die Fähigkeit beziehungsweise den Versuch eines Individuums, sich in andere hineinzusetzen, um deren Wahrnehmungen, Gedanken und Absichten zu verstehen“ (Förstl, 2007, S.4) ToM bildet die Basis des sozialen Verhaltens, da sie es ermöglicht ein Gefühl und ein gewisses Gespür für die Gedanken und Bedürfnisse anderer zu entwickeln. Sie ist essentiell für die Nachvollziehbarkeit des Verhaltens und der Intentionen anderer Personen sowie für ein differenziertes Verständnis ihrer Perspektiven. (Förstl, 2007).

Es handelt sich bei der ToM um alltagspsychologische Konzepte, die man einsetzt, um sich selbst und den anderen mentale Zustände zuzuschreiben (was wir wissen, wollen, denken und fühlen) (Sodian, 2007)

Evelyn C. Ferstl (2007, S.68) versteht unter der ToM „die grundlegende Fähigkeit sich in andere hineinzusetzen, ihre Gefühle, ihr Denken und Wissen

einzubeziehungen und zu erkennen, dass die Handlungen anderer Menschen durch ihre mentalen und emotionalen Zustände bedingt sind.“

Zwei ganz wesentliche Aspekte, die durch die ToM ermöglicht werden, sind die Erklärung und Interpretation sowie die Vorhersage menschlichen Verhaltens. In diesem Zusammenhang spielen drei grundlegende Konzepte eine Rolle: die *Überzeugung*, das *Bedürfnis* und die *Handlung*. Diese Konzepte bilden die wesentliche Voraussetzung und den Grundstein für die Erklärung und Vorhersage menschlichen Verhaltens. Durch die Überzeugungen, die uns die Beschaffenheit der Welt oder das wofür wir sie halten widerspiegeln und die Bedürfnisse, die Ausdruck dessen sind, wie wir wünschen, dass die Welt ist, entstehen Handlungen (Astington, 2000). Alle drei Elemente beeinflussen sich wechselseitig.

Weitere ontogenetische Voraussetzungen für die ToM sind die Ausbildung eines Selbstkonzeptes, die Fähigkeit zur Empathie und das Verständnis für die Zeit. Bischof-Köhler (2000) konnte im Zuge ihrer Experimente diesen Zusammenhang nachweisen und zusätzlich feststellen, dass sich diese Kompetenzen im Laufe des vierten Lebensjahres ausbilden (Ploog, 2007).

4.3. Konzept der falschen Überzeugungen – False belief

Wie bereits im vorigen Absatz erwähnt wurde, ist für die Überprüfung der ToM die Repräsentation einer falschen Überzeugung einer Person über eine Situation erforderlich. Wimmer und Perner (1983) orientierten sich an Premacks und Woodruffs Forschungsergebnissen und entwickelten infolgedessen ein experimentelles Paradigma zur Untersuchung des Verständnisses falscher Überzeugungen bei Kindern. Sie waren die ersten, die eine systematische Untersuchung zum kindlichen Verständnis falschen Glaubens durchführten. Wimmer und Perner (1983) konstruierten die klassische *false-belief*-Aufgabe, die sogenannte *Maxi-Aufgabe*, bei der ein Versuchsleiter der Versuchsperson eine Geschichte mit Figuren und Objekten vorspielt. Die Geschichte lautet folgendermaßen: „Ein Bub (Maxi) legt Schokolade an einen bestimmten Ort X (in den Schrank). Maxi geht daraufhin auf den Spielplatz. In der Zwischenzeit kommt die Mutter und legt die Schokolade von dem Ort X, dem Schrank in den Ort Y (in

die Schublade). Daraufhin verlässt die Mutter die Szene und Maxi kommt zurück. Die Frage ist nun: *Wo wird Maxi die Schokolade suchen?*“ (Wimmer & Perner, 1983)

Die Ergebnisse der Autoren zeigen, dass fast alle Kinder unter 3 Jahren diese Aufgabe falsch beantworten, indem sie davon ausgehen, dass Maxi die Schokolade in der Schublade (Ort Y) suchen wird, wo sie sich auch tatsächlich befindet. Kinder unter drei Jahren verfügen noch nicht über die Fähigkeit der Unterscheidung der eigenen Überzeugung von der Überzeugung anderer. In Bezug auf die Maxi-Aufgabe bedeutet das, dass das Kind realisieren muss, dass Maxi über andere Überzeugungen verfügt als es selbst und dass Maxis falsche Überzeugungen ihn auch zu einem falschen Handeln bewegen (Astington & Dack, 2008).

Um nun ein Verständnis von mentalen Zuständen und deren korrekter Zuschreibung zu erlangen und zu gewährleisten, muss das Kind erkennen, dass sich diese mentalen Zustände und Überzeugungen in der Realität abbilden und in weiterer Folge im Verhalten manifestieren können. Es handelt sich dabei jedoch um mentale und internale Zustände, die sich von der Realität unterscheiden können. Die Grundlage für das kindliche Verständnis falscher Überzeugungen anderer Personen, die im Widerspruch zum realen Geschehen stehen, bildet die Unterscheidung zwischen den eigenen Überzeugungen und Repräsentationen von Zuständen und denen der Welt. (Wellman, Cross & Watson, 2001)

Bei der angeführten Maxi-Aufgabe von Wimmer und Perner (1983) handelt es sich um eine *change-of.location*-Aufgabe, die an das Kind die Anforderung des Geschichtenverstehens stellt. Eine weitere häufig eingesetzte *false-belief*-Aufgabe, die diese Fertigkeit nicht voraussetzt ist die *Smarties*-Aufgabe. Hierbei handelt es sich um eine *unexpected-identity*-Aufgabe, bei der das Kind selbst die Erfahrung macht, einen falschen Glauben und falsche Überzeugungen zu haben. Dem Kind wird eine bekannte Verpackung (z.B. Smarties-Schachtel) gezeigt, und es wird nach dem von ihm erwarteten Inhalt gefragt. In der Verpackung befinden sich jedoch nicht die erwarteten Smarties, sondern ein anderer Inhalt, ein Bleistift. Es folgt eine zweite Frage, worauf das Kind antworten muss, was ein anderes

Kind glauben bzw. sagen würde, wenn es den tatsächlichen Inhalt (Bleistift) der Verpackung nicht gesehen hätte (Hogrefe, Wimmer & Perner, 1986).

Es gibt noch weitere Varianten der false-belief-Aufgaben, auf die hier jedoch nicht mehr näher eingegangen wird (Milligan, Astington & Dack, 2007). Die angeführten Darstellungen beschränken sich auf die in der Forschung am häufigsten verwendeten false-belief-Aufgaben, die Maxi- und die Smarties-Aufgabe.

Abschließend ist nochmals zu erwähnen, dass im Zuge vieler Studien gezeigt werden konnte, dass das Kind für die korrekte Lösung der false-belief-Aufgaben ein Verständnis entwickelt haben muss, dass sich einerseits die eigenen Überzeugungen von denen der anderen unterscheiden können und dass andererseits die Beziehung zwischen Person und Umwelt durch deren mentale Repräsentationen entstehen (Astington & Dack, 2007).

4.4. Entwicklung der ToM

Die Entwicklung der ToM setzt den Erwerb einer Reihe von Fähigkeiten voraus. Wie bereits erwähnt, konnte in einer Vielzahl von Studien (Wimmer & Perner, 1983; Wellman, Cross & Watson, 2001; Astington & Dack, 2008) gezeigt werden, dass Kinder erst im Altersbereich zwischen etwa drei und vier Jahren, sowohl das Konzept der Überzeugung erwerben als auch die begrifflichen Differenzierungen zwischen Aussehen und Realität und Überzeugung und Realität vornehmen können. Kinder unter etwa dreieinhalb Jahren weisen erhebliche begriffliche Probleme auf, in dem sie scheinbar nicht über den Begriff der Überzeugung und über das grundlegende Verständnis verfügen, dass subjektive Überzeugungen und die Repräsentation der realen Welt voneinander differieren können (Sodian, 2002). Für die Entwicklung und Ausbildung der ToM ist somit nicht nur das Bewusstsein mentaler Zustände und die Fähigkeit deren Zuschreibung sowie die Verbindung von Bedürfnissen und Handlungen, sondern auch das Verständnis für richtige und falsche Überzeugungen erforderlich. Grundstein für dieses Verständnis, das die Vorhersage von Verhalten ermöglicht, ist die Entwicklung eines Verständnisses für Metarepräsentationen. Darunter

versteht man das grundlegende Verständnis dafür, dass es sich bei den Überzeugungen, Bedürfnissen und Absichten von Personen um mentale Repräsentationen handelt, die deren Handlungen steuern (Astington & Dack, 2008).

Ab dem Alter von vier Jahren zeigt sich ein gefestigtes Verständnis von den Kernbegriffen *Absicht* und *Überzeugung*. Kinder dieses Alters verstehen, dass eine Person falsche Überzeugungen über einen Sachverhalt haben kann und dementsprechende Handlungen setzt. Des Weiteren können Vierjährige Wissen durch den visuellen Zugang, durch das Sehen gewisser Objekte und Handlungen, erfassen. Hinsichtlich der Einsicht in den eigenen Lernprozess weisen Kinder dieser Altersgruppe jedoch noch Defizite auf, da die Unterscheidung zwischen aktuellen Lernereignissen und Vorwissen noch nicht möglich ist. Diese Differenzierung kann erst im Alter von fünf bis sechs Jahren vorgenommen werden. Bis zum Alter von sechs Jahren kommt es im Rahmen der ToM-Entwicklung noch zu wesentlichen mentalen, begrifflichen Erweiterungen und Differenzierungen. Im Zuge einiger Untersuchungen konnte gezeigt werden, dass Sechsjährige verstehen, dass eine Überzeugung über eine Überzeugung einer anderen Person falsch sein kann (Perner & Wimmer, 1986) und dass auch schlussfolgendes Denken zu Wissen führt (Sodian & Wimmer, 1987, zitiert nach Sodian, 2002).

4.4.1 Entwicklung einer fortgeschrittenen ToM

Im Grundschulalter kann eine sich eine zunehmende Beweglichkeit und Flexibilität im Denken beobachtet werden (Sodian, 2002). Ein wesentliches Merkmal in der Entwicklung der ToM bei Kindern, im Alter von sieben bis acht Jahren ist, dass sie eine Vorstellung von einer kontinuierlichen gedanklichen Aktivität im Sinne eines *Bewusstseinsstromes* entwickeln. Kinder diesen Alters haben eine Vorstellung einer *inneren Sprache* entwickelt und begreifen, dass Gedanken unwillkürlich und oft schwer zu unterdrücken sind (Flavell et al., 1997, Sodian, 2002)

Es zeigt sich, dass Kinder im Grundschulalter bereits über die Fähigkeit der Zuschreibung mentaler Zustände zweiter und höherer Ordnung verfügen, was bedeutet, dass sich die Zuschreibung mentaler Zustände nicht mehr nur auf Repräsentationen der Realität beschränken, sondern auch Repräsentationen mentaler Repräsentationen der Realität miteinschließen.(z.B.“ Maxi glaubt, dass Susi glaubt, der Eisverkäufer sei im Park) Eine ToM zweiter Ordnung bildet die Grundlage für das Verständnis von Ironie und Witz sowie für den Einsatz und das Wissen von Metaphern (Sodian, 2002).

Darüber hinaus verstehen Kinder dieses Alters, dass Erwartungen und Vorurteile die Interpretation aktueller Ereignisse beeinflussen können. Sie beginnen eine Ahnung von der Wirkung von Vorurteilen und Voreingenommenheiten zu entwickeln und sind fähig dazu Personen, psychologische Eigenschaften zuzuschreiben (Sodian, 2002).

4.5. Erklärungsansätze und Theorien zur ToM

In diesem Abschnitt wird ein kurzer Überblick über unterschiedliche theoretische Ansätze zur Erklärung der ToM gegeben.

4.5.1 Theorie-Theorie

Bei der Theorie-Theorie handelt es sich um eine *intuitive* Theorie, da die mentalen Zustände nicht durch die direkte Beobachtung zugänglich sind, sondern vielmehr wie theoretische Terme erschlossen werden müssen (Sodian, 2002). Im Rahmen dieses Konzepts werden die Entwicklungsveränderungen und-fortschritte des begrifflichen Systems im Sinne eines *Theoriewandels* verstanden. Die Bezeichnung *Theoriewandel* bezieht sich auf den Prozess der kontinuierlich zunehmenden gedanklichen und geistigen Beweglichkeit. Grundgedanke dieses Ansatzes ist ein theorieähnliches System begrifflichen Wissens, das für die Interpretation und das Verständnis des eigenen und fremden Verhaltens verantwortlich ist.

Wie bereits zuvor erwähnt verfügen Kinder nicht von Beginn an über das Konzept der Überzeugung. Sie erklären sich zunächst das Verhalten anderer Personen in den Begriffen von Wünschen, Bedürfnissen und Emotionen. Im weiteren Verlauf vollzieht sich der Wandel zu einer Theorie, einer theoretischen Grundlage, in der das Konzept der Überzeugung verankert und maßgeblich für das Verständnis und die Nachvollziehbarkeit menschlichen Verhaltens ist. Das ursprüngliche nichtrepräsentationale Begreifen und Verstehen wächst zu einem repräsentationalen Verständnis des mentalen Bereichs heran (Gopnik & Wellman, 1994, zitiert nach Sodian, 2007). Diese repräsentationale ToM kann als echte kausale Theorie mentaler Vorgänge angesehen werden. Im Rahmen der Theorie-Theorie wird davon ausgegangen, dass sich das Verständnis der eigenen und der fremden mentalen Zustände ziemlich gleichzeitig ausbildet (Sodian, 2007).

Die derzeitige empirische Befundlage spricht für die Geltung und Annahme der Theorie-Theorie, da bei Kindern, wie vorher bereits angeführt, die Konzeptualisierung eigener und fremder mentaler Zustände ziemlich simultan verläuft (Sodian, 2002).

4.5.2 Simulationstheorie

Simulationstheorien hingegen, verfolgen den Ansatz, dass unsere alltagspsychologischen Interpretationen und Konstrukte durch die unmittelbaren Erfahrungen des eigenen psychischen Geschehens entstehen. Die Simulationsannahme geht davon aus, dass Kinder zwar einen unmittelbaren Zugang zu ihrem eigenen mentalen und emotionalen Geschehen haben, jedoch Probleme in Hinblick auf das Verständnis und die Nachvollziehbarkeit der mentalen Vorgänge anderer Personen aufweisen (Sodian, 2007). Demzufolge geht es um die Projektion eigener mentaler Zustände auf jemanden anderen, wenn die mentalen Zustände anderer simuliert werden sollen. Das bedeutet nach Harris (1994), dass die Simulation des mentalen Zustands anderer, der sich von dem eigenen unterscheidet, die Unterdrückung und das Ignorieren der eigenen mentalen Zustände voraussetzt und erfordert, um den Zustand des anderen unter

Berücksichtigung der für ihn relevanten Bedingungen zu simulieren (Sodian, 2007).

4.5.3 Modularitätstheorie

Im Rahmen der Modularitätstheorien wird ein metarepräsentationales Verständnis falscher Überzeugungen postuliert, das bereits vor der Fähigkeit, Aufgaben zum Verständnis falscher Überzeugungen zu lösen, gegeben ist. Die Entwicklung und Ausbildung der ToM wird auf die neurologische Reifung von drei bereichsspezifischen modularen Mechanismen zur Repräsentation der Eigenschaften von Agenten zurückgeführt.

- **ToBy (Theory-of-Body-Mechanismus)**

Dieser Mechanismus bildet sich in der ersten Hälfte des ersten Lebensjahres aus und befähigt den Säugling durch spontane Bewegungsfähigkeit Agenten von Nichtagenten zu unterscheiden. Aufbauend auf diesem Mechanismus greifen zwei weitere fortgeschrittene Mechanismen:

- **TOMM1**

Gegen Ende des ersten Lebensjahres sind Repräsentationen intentionaler Agenten möglich, deren Handeln hinsichtlich konkreter Ziele zu interpretieren ist.

- **TOMM2**

Dieser Mechanismus beginnt im Alter von etwa 18 Monaten zu greifen und führt zur Ausbildung der metarepräsentationalen Fähigkeit, die relevant für die Repräsentation der propositionalen Einstellungen von Agenten (z.B. glauben, dass-wünschen, dass) sind (Leslie, 1994; Sodian, 2007).

Vertreter der nativistischen Modularitätstheorien wie Fodor (1992) und Leslie (1994) gehen davon aus, dass die Zuschreibung von Absichten und Überzeugungen eine im Zuge der Evolution entstandene und ausgebildete spezifisch menschliche Fähigkeit ist, die bereits in der frühen Kindheit vorhanden ist. Des Weiteren machen sie die Gedächtnis- und

Aufmerksamkeitsanforderungen der Aufgaben für das Scheitern jüngerer Kinder in false-belief-Aufgaben verantwortlich und zweifeln echte und nachhaltige Veränderungen im Verständnis des mentalen Bereichs in der Kindheit an (Fodor, 1992; Leslie, 1994; Sodian, 2002).

4.6. Geschlechtsspezifische Unterschiede

Im Folgenden wird anhand einiger Studien der Aspekt geschlechtsspezifischer Unterschiede in der ToM diskutiert. Im Zuge der Durchsicht der Literatur zeigte sich, dass bei dem Großteil der Untersuchungen der Fokus auf jüngeren Kindern im Kleinkind-bis Grundschulalter liegt. Des Weiteren zeigte sich, dass in den meisten Studien, die sich mit individuellen Unterschieden im ToM-Verständnis beschäftigen, nicht explizit der Geschlechtsaspekt beleuchtet und berücksichtigt wird. Es gibt diesbezüglich nur wenige Untersuchungen und Forschungsergebnisse, die die Thematik der individuellen Unterschiede im Bereich der sozial-kognitiven Funktionen unter dem Gesichtspunkt des Geschlechts und der geschlechtsspezifischen Unterschiede betrachten. In der Literatur demonstrieren Forschungsergebnisse geschlechtsspezifische Unterschiede in der ToM , insbesondere in vier Bereichen:

1. Verständnis falscher Überzeugungen
2. Soziales Verständnis, soziale Kompetenz und Empathiefähigkeit
3. Sprachliche Fähigkeiten
4. Autismus

Der dritte Punkt der sprachlichen Fähigkeiten wird in dieser Arbeit nicht als extra Punkt behandelt, jedoch in die Darstellungen der anderen Punkte mit einbezogen.

4.6.1 Unterschiede im Verständnis falscher Überzeugungen

Cutting und Dunn (1999) untersuchten Kinder im Alter von vier Jahren in Hinblick auf individuelle Unterschiede in ihrer kindlichen, sozialen Kognition. Hauptaugenmerk der Untersuchung lag auf der Erfassung von Unterschieden sowohl im Verständnis falscher Überzeugungen als auch im emotionalen

Verständnis. Darüber hinaus sollten mögliche Zusammenhänge mit sprachlichen Fähigkeiten sowie mit dem familiären Background untersucht werden. Zusätzlich wurde auch der Aspekt des Geschlechts und diesbezügliche geschlechtsspezifische Unterschiede berücksichtigt. Die Ergebnisse veranschaulichen, dass der kindliche familiäre Background einen starken Einfluss auf die Entwicklung der sozialen Kognition, insbesondere auf das Verständnis falscher Überzeugungen hat (Cutting & Dunn, 1999). Hinsichtlich der Bedeutung sprachlicher Fähigkeiten im Zusammenhang mit der ToM konnten die bestehenden Befunde in der Literatur belegt werden (Jenkins & Astington, 1996, Astington & Jenkins, 1999; Charman et al., 2002). Es konnte demonstriert werden, dass die Sprache nachweislich eine sehr wichtige Rolle im Rahmen der ToM-Entwicklung hat. Cutting und Dunn heben in ihrer Studie diesbezüglich zwei Aspekte der Sprache heraus, die besonders zur Entwicklung der sozialen Kognition beitragen: der rezeptive Wortschatz und der Ausdruck der Sprache. Die Sprache leistet einen grundlegenden Beitrag zum ToM-Verständnis, insbesondere in Bezug auf das Verständnis falscher Überzeugungen. Darüber hinaus konnte ein Zusammenhang zwischen der Leistung in den false-belief-Aufgaben und dem emotionalen Verständnis festgestellt werden. In Bezug auf Geschlechtsaspekt präsentierten Cutting und Dunn (1999), dass Mädchen bessere Leistungen in den false-belief-Aufgaben erbrachten als Burschen. Des Weiteren zeigte sich ein schwacher, aber nicht signifikanter Vorteil der Mädchen im emotionalen Verständnis. Im Bereich der sprachlichen Fähigkeit konnte zwar ein Unterschied, jedoch kein signifikanter geschlechtsspezifischer Unterschied festgestellt werden (Bosacki et al., 1999; Cutting & Dunn, 1999). Es zeigte sich, dass bei den Burschen alle Aspekte der Sprache in Verbindung mit der sozialen Kognition standen. Bei den Mädchen hingegen erwiesen sich narrative, expressive Fähigkeiten als besonders relevant bezüglich der sozialen Kognition (Cutting & Dunn, 1999).

Eine weitere Studie, die sich mit der Erforschung und Erfassung geschlechtsspezifischer Unterschiede im Verständnis falscher Überzeugungen beschäftigt, wurde von Charman, Ruffman und Clements (2002) durchgeführt. Sie untersuchten Kinder im Alter von etwa zwei bis sechs Jahren und fanden einen

leichten Vorteil der Mädchen in den false-belief-Ausgaben gegenüber den Burschen. Dieser Vorteil zeigte sich jedoch nur bei den jüngeren Mädchen.

4.6.2 Unterschiede in der sozialen Kompetenz

In einer Studie von Walker (2005) wurde die Beziehung zwischen dem ToM – Verständnis bei Vorschulkindern, im Alter von drei bis fünf Jahren, ihrem Sozialverhalten und ihrer sozialen Kompetenz gegenüber Gleichaltrigen untersucht. Den Kindern wurden zwei unterschiedliche ToM-Aufgaben zur Erfassung des Verständnisses falscher Überzeugungen vorgegeben. Des Weiteren wurden die Einschätzungen der LehrerInnen in Bezug auf das Sozialverhalten der Kinder erfasst. In Bezug auf das Alter zeigte sich, dass ältere Kinder bessere Leistungen in den false-belief Aufgaben erzielten als jüngere. Die Ergebnisse bezüglich der sozialen Kompetenz belegen, dass die Fähigkeit der ToM eine Vorhersage des sozialen Verhaltens und der sozialen Kompetenz erlaubt. Diesbezüglich zeigen sich klare geschlechtsspezifische Unterschiede. Bei den Burschen konnte ein positiver Zusammenhang zwischen dem ToM-Verständnis und aggressivem und störenden Verhalten festgestellt werden, während bei Mädchen ein positiver Zusammenhang für prosoziales Verhalten beobachtet werden konnte. Die Ergebnisse von Charman et al., (2002), in Bezug auf den leichten Vorteil der Mädchen in den false-belief-Aufgaben, konnten repliziert werden (Walker, 2005).

Bosacki, University und Astington (1999) untersuchten im Rahmen ihrer Studie Mädchen und Burschen in der Präadoleszenz im Alter von 10 bis 13 Jahren. Hauptaugenmerk ihrer Studie lag auf der Erforschung der ToM bei Kindern in der Präadoleszenz, hinsichtlich individueller Unterschiede im ToM-Verständnis, als Teil des sozialen Verständnisses. Es sollte überprüft werden, ob ein Zusammenhang zwischen dem ToM-Verständnis und der sozialen Kompetenz in der Präadoleszenz besteht und ob sich dieser für Mädchen und Burschen unterschiedlich gestaltet. Darüber hinaus wurde auch der Wortschatz erhoben (Bosacki et al., 1999). Es konnte ein Zusammenhang zwischen dem sozialen

Verständnis (ToM) und der Peer-bezogenen sozialen Interaktionsfähigkeit sowie deren Wortschatz festgestellt werden. Bezüglich der Untersuchung der geschlechtsspezifischen Unterschiede konnten signifikante Ergebnisse nachgewiesen werden. Demnach erreichen Mädchen signifikant höhere Ergebnisse im Bereich des sozialen Verständnisses (ToM) und der sozialen Kompetenz als Burschen. Mädchen wurden von Gleichaltrigen als sozialer und liebenswerter eingeschätzt. Überraschenderweise stellten Bosacki et al. (1999) keine geschlechtsspezifischen Unterschiede in der Wortschatz-Aufgabe fest. Diese Ergebnisse sind nicht konsistent mit anderen Forschungsergebnissen, die für Mädchen einen klaren Vorteil im Bereich des Wortschatzes und der sprachlichen Fähigkeit deklarieren (Happé, 1995; Astington & Jenkins, 1999; Cutting & Dunn, 1999). Des Weiteren zeigte sich, dass das Geschlecht des Lehrers eine gewisse Rolle im Rahmen des kindlichen sozialen Verständnisses spielt. Mädchen wiesen bei männlichen Lehrern die höchsten Werte in der empathischen Sensibilität auf und Burschen erreichten bei weiblichen Lehrerinnen die niedrigsten Werte bezüglich der empathischen Sensibilität. Zusammenfassend kann man sagen, dass die Ergebnisse dieser Studie die Annahme unterstützen, dass die Wahrnehmungen und die Erwartungen der LehrerInnen und Gleichaltrigen über die/den jeweilige/n SchülerIn von traditionellen sozialen und gesellschaftlichen geschlechtsrollenspezifischen Stereotypen geleitet werden und einen Effekt auf deren kognitive und soziale Fähigkeiten haben.

O'Connor und Hirsch (1999) konzentrierten sich in ihrer Studie auf die Untersuchung intraindividuelle Unterschiede in der *Mentalisierung* in der frühen Adoleszenz bei Burschen und Mädchen im Alter von 12 bis 14 Jahren. Der Begriff der *Mentalisierung* steht sowohl für das Verständnis der eigenen Gedanken und Gefühle als auch für das der anderen (Fonagy et al., 1991). Der Fokus dieser Untersuchung lag auf der Erfassung intraindividuelle Unterschiede in der Mentalisierungsfähigkeit im Beziehungskontext mit sich selbst und mit den anderen (O'Connor & Hirsch, 1999). Eines der Ziele bestand darin, zu untersuchen, ob diese Fähigkeit der Mentalisierung auf einem gewissen Beziehungskontext und einer Beziehungsspezifität basiert oder ob es sich diesbezüglich um eine für das Individuum typische Charakteristik handelt. Bei

den Resultaten zeigte sich bereits zu Beginn kein signifikantes Ergebnis für das Geschlecht. Infolgedessen wurden geschlechtsspezifische Unterschiede im Rahmen dieser Untersuchung nicht mehr bei den weiteren Berechnungen berücksichtigt. Die Endergebnisse sprechen für intraindividuelle Unterschiede in der Mentalisierungsfähigkeit in der frühen Adoleszenz und weisen auf eine Verbindung zwischen dem Grad der Mentalisierung und der affektiven Qualität der Beziehung auf.

4.6.3 ToM und Autismus

In Bezug auf relevante Forschungsergebnisse, die den Bereich der ToM im Zusammenhang mit geschlechtsspezifischen Unterschieden untersuchen, sind unbedingt auch die interessanten und bahnbrechenden Studien von Simon Baron-Cohen zu nennen. Baron-Cohen, ein britischer Psychologe und Direktor des Autismus-Forschungszentrums in Cambridge, konzentriert sich in seinen Untersuchungen auf die Erforschung der tiefgreifenden Entwicklungsstörung des Autismus und entwickelte zwei wegweisende Theorien über seine Entstehung, die *extreme male brain-theory* und die *empathising-systemising-theory*. Die Bedeutung geschlechtsspezifischer Unterschiede wird in seinen Theorien und Untersuchungen ebenfalls thematisiert (Baron-Cohen & Hammer, 1997; Baron-Cohen et al., 2004; 2005). Zum besseren Verständnis wird an dieser Stelle kurz der Begriff des Autismus erläutert. Autismus ist eine tiefgreifende Entwicklungsstörung, die mit gravierenden Defiziten und Beeinträchtigungen in der sozialen Entwicklung einhergeht und durch repetitive, stereotype Verhaltensweisen und einer sehr begrenzten Vorstellungskraft charakterisiert ist (American Psychiatric Association, 1994).

Personen mit klassischem Autismus, HFA (High-functioning Autismus = Kriterien für Autismus werden erfüllt bei dem Vorhandensein eines normalen IQ; World Health Organization, 1994) und AS (Asperger Syndrom= Kriterien für Autismus werden ebenfalls erfüllt, aber ohne Anzeichen einer Entwicklungsverzögerung im kognitiven und sprachlichen Bereich; World Health organization, 1994) weisen starke Defizite und Beeinträchtigungen in der Tom-

Entwicklung auf, was auch *als Grade der mind-blindness* bezeichnet wird (Baron-Cohen, 1995; Baron-Cohen, Knickmeyer und Belmonte, 2005).

Baron-Cohen befasste sich im Zusammenhang mit der Erforschung von Autismus mit der Untersuchung von geschlechtsspezifischen Unterschieden in der Gehirnentwicklung und kam zu dem Schluss, dass die Gehirne von Frauen und Männern bereits im Mutterleib unterschiedlich vorprogrammiert werden. Demzufolge bezeichnet er das männliche Gehirn als das *S-Gehirn* und das weibliche Gehirn als das *E-Gehirn*. Das *S-Gehirn* steht im Rahmen der *Empathizing-Systemizing – Theorie* (E-S-Theorie) im Zusammenhang mit dem Begriff des *Systematisierens*, worunter man die Analyse von Systemen, gemäß ihren zugrundeliegenden Regeln und die Vorhersage ihres Verhaltens versteht. Das *E-Gehirn* steht in Verbindung mit dem Begriff des *Empathizing*, des Einfühlens der Empathie, worunter die Fähigkeit verstanden wird, die Gefühle, die Gedanken und das Verhalten anderer zu identifizieren, nachzuvollziehen und vorhersagen zu können, um in weiterer Folge adäquat darauf zu reagieren (Baron-Cohen, Knickmeyer & Belmonte, 2005). Die E-S-Theorie besagt nun, dass geschlechtsspezifische Unterschiede durch die unterschiedliche Ausprägung in den zwei Dimensionen des Systematisierens und des Einfühlens definiert und veranschaulicht werden. Demnach weisen Männer eine ausgeprägte und stärkere Neigung zum Systematisieren und Frauen eine klare Überlegenheit in der Empathiefähigkeit auf (Baron-Cohen & Wheelwright, 2004; Baron-Cohen et al., 2005). Die Tatsache, dass biologische Faktoren das Verhalten und das soziale Interesse eines Menschen prägen und modellieren, konnte eindrucksvoll in Experimenten mit einen Tag alten Säuglingen demonstriert werden. Es wurde männlichen und weiblichen Säuglingen entweder ein lachendes Gesicht oder ein mechanisches Mobile gezeigt. Die Resultate zeigten, dass Mädchen durchschnittlich mehr Zeit damit verbrachten das lachende Gesicht anzuschauen und Burschen demgegenüber mehr Zeit mit der Fixierung des Mobiles verbrachten (Baron-Cohen et al., 2005). In diesem Zusammenhang ist der Aspekt und Einfluss des Testosteronspiegels anzuführen. Baron-Cohen untersuchte im Rahmen eines Experiments schwangere Frauen und erhob den Testosteronspiegel ihrer ungeborenen Babies. Es konnte nachgewiesen werden, dass Kinder, die im

Mutterleib einer hohen Testosteronkonzentration ausgesetzt waren, gegenüber gesunden Kindern, signifikant weniger Blickkontakt suchten, einen kleineren, aber spezifischeren und qualitativ ausgereifteren Wortschatz und eine verzögerte soziale Entwicklung aufwiesen. In Bezug auf Cohen's E-S-Theorie ist dieser Sachverhalt so zu verstehen, dass sich die Gehirne dieser Kinder, aufgrund des erhöhten Testosteronspiegels im Mutterleib, dahingehend entwickeln, dass sie besonders gut Muster erkennen und Systeme analysieren können (Baron-Cohen, 2002; Baron-Cohen & Hammer, 1997).

In einer seiner Studien untersuchten Baron-Cohen und Wheelwright (2004) sowohl die Empathiefähigkeit bei Erwachsenen mit AS (Asperger Syndrome) und HFA (High functioning Autism) als auch diesbezügliche geschlechtsspezifische Unterschiede. An dieser Stelle soll, zum besseren Verständnis, nochmals kurz auf den Begriff der Empathie eingegangen werden.

Empathie ist eine wichtige, sowohl kognitive als auch affektive Kernfähigkeit der ToM und beschreibt die Fähigkeit sich in andere Personen, in ihre Gedanken- und Gefühlswelt hineinversetzen und ihr Verhalten vorhersagen zu können. Diese Empathiefähigkeit ist sehr defizitär ausgeprägt oder kaum vorhanden bei Personen mit Autismus (Baron-Cohen & Wheelwright, 2004). Die Empathiefähigkeit wurde im Rahmen der Untersuchung von Baron-Cohen & Wheelwright (2004) mittels eines selbstkonstruierten Questionnaire dem *Empathy-Quotient* (EQ) gemessen. Viele der Items des EQ stehen für die Notwendigkeit einer entwickelten ToM und für dessen Beantwortung und Verständnis eine intakte ToM Voraussetzung ist. Die Resultate demonstrieren, wie vermutet, dass Personen mit AS/HFA signifikant niedrigere Werte im EQ erzielen als die Personen der Kontrollgruppe. Diese Ergebnisse stützen die Annahme, die bereits in vorherigen Studien (Baron-Cohen, 1995) bestätigt werden konnte, dass die Empathiefähigkeit und somit auch die ToM bei Personen mit Autismus defizitär ausgeprägt ist. In Bezug auf den Aspekt geschlechtsspezifischer Unterschiede in der Empathiefähigkeit konnten die Ergebnisse vorheriger Studien repliziert werden, die eine weibliche Überlegenheit in der Empathie nachweisen (Davis & Franzoi, 1991; Baron-Cohen & Wheelwright, 2004). Die Resultate dieser Untersuchung sind ebenfalls

konsistent mit der bereits erwähnten *Extreme male brain theory* (EMB-Theorie, Baron-Cohen, 2002; 2003; Baron-Cohen & Hammer, 1997), die von zwei Dimensionen, dem *Systemizing* (Systematisieren) und dem *Empathizing* (Einfühlens, Empathiefähigkeit) ausgeht und besagt, dass Personen mit Autismus ein überdurchschnittlich ausgeprägtes männliches Profil des Gehirns aufweisen, was auf die erhöhte Testosteronkonzentration im Mutterleib zurückzuführen ist (Baron-Cohen & Hammer, 1997; Barn-Cohen, 2002; 2003). Des Weiteren erklärt die EMB-Theorie, dass Personen mit Störungen des autistischen Formenkreises charakteristische Beeinträchtigungen und Schädigungen im Bereich der Empathie und im Gegensatz dazu, eine intakte und oftmals überdurchschnittlich ausgeprägte Systemisierungsfähigkeit aufweisen (Baron-Cohen & Wheelwright, 2005).

In diesem Zusammenhang ist eine sehr interessante Studie von Baron-Cohen et al. (1998) zu nennen, in der die Annahme untersucht wurde, dass Autismus vergleichsweise häufiger in Familien auftritt, in denen die Eltern und/oder Verwandten ersten Grades in den Bereichen Physik, technische Konstruktion und Mathematik tätig sind und über stärker ausgeprägtes naturwissenschaftliches Verständnis in Alltagssituationen, im Gegensatz zu *alltagspsychologischen Konzepten* verfügen (Baron-Cohen, Bolton, Wheelwright, Scahill, Short, Mead & Smith, 1998). Baron-Cohen et al. (1998) gingen im Rahmen ihrer Untersuchung von der Annahme aus, dass Autismus mit einem defizitären alltagspsychologischen Verständnis, beziehungsweise einem sozialen Verständnis und einem intakten und überlegenen naturwissenschaftlichen Verständnis, dem Verstehen unbelebter Objekte und physikalischer Vorgänge und Phänomene, einhergeht. Im Zuge der Studie wurden StudentInnen, die eines der drei Fächer (Physik, technische Konstruktion bzw. Ingenieurwesen, Mathematik) studierten, mit StudentInnen, die englische bzw. französische Literatur studierten, in Bezug auf das familiäre Vorkommen von Autismus untersucht. Die Ergebnisse bestätigen die Annahme, dass ein signifikanter Zusammenhang zwischen Personen mit einer ausgeprägten und überlegenen naturwissenschaftlichen Haltung und dem familiären Auftreten von Autismus beobachtet werden kann (Baron-Cohen et al., 1998).

Abschließend ist zu sagen, dass die vorgestellten Studien einen differenzierten Einblick in die Thematik geschlechtsspezifischer Unterschiede in der ToM liefern und mehrfach eine weibliche Überlegenheit in der ToM und somit auch in der Empathiefähigkeit belegen. Die Experimente und Studien von Baron-Cohen sind diesbezüglich besonders bedeutend, da auf biologische Einflussfaktoren und die vorgeburtliche Prägung des Gehirns im Zusammenhang mit geschlechtsspezifischen Unterschieden eingegangen wird. Es wird sehr verständlich veranschaulicht, dass sich Männer und Frauen in kognitiven Bereichen aufgrund ihrer Entwicklung unterscheiden und im Zuge dessen, unterschiedliche Schwerpunkte in ihrem Denken und in ihrer Wahrnehmung ausbilden. Dies spiegelt sich in der Tatsache wider, dass Männer bevorzugt in Systemen denken, sich durch einen ausgeprägten technischen Sachverstand auszeichnen und mit der Funktionalität von Objekten beschäftigen. Frauen hingegen begreifen, erfassen und erklären sich die Welt mit Hilfe der Empathie und haben somit einen anderen Zugang zu Problemlöse- und Bewältigungsstrategien. In diesem Zusammenhang ist wichtig anzumerken, dass es bei der Darstellung und Auseinandersetzung mit den genannten geschlechtsspezifischen Unterschieden nicht um eine Bewertung von besser oder schlechter geht. Es soll gezeigt und beleuchtet werden, dass es zwar Unterschiede zwischen Männer und Frauen in kognitiven Bereichen gibt, aber dass eben diese Unterschiede auch mögliche Erklärungen für die unterschiedlichen kognitiven Leistungen, Bewältigungsstrategien und Herangehensweisen von Frauen und Männern darstellen.

4.7. Wechselwirkungen und Zusammenhänge

Im Folgenden soll auf mögliche und erwartete Wechselwirkungen und Zusammenhänge der drei Konstrukte *Theory of Mind*, *Stereotype Threat* und *Risikoverhalten* eingegangen und ein Bezug zu der Fragestellung dieser Arbeit hergestellt werden. In dieser Arbeit wird, im Rahmen der Bewältigung EMS-ähnlicher Aufgaben, die *Theory of Mind* (bzw. das erfolgreiche Bearbeiten der *Theory of Mind*-Aufgaben) im Zusammenhang mit dem Phänomen *Stereotype*

Threat und Risikoverhalten, unter Berücksichtigung geschlechtsspezifischer Unterschiede bei weiblichen und männlichen Maturanten untersucht.

Die Auseinandersetzung mit der Fachliteratur zeigte, dass diese drei Konstrukte bis dato noch nicht in diesem Zusammenhang untersucht worden sind und somit eine neue und bis jetzt noch nicht untersuchte Fragestellung in diesem Bereich darstellt.

Im Zusammenhang mit dem Konstrukt *Stereotype Threat* gibt es zahlreiche Studien (Steele et al., 1995; Steele, 1997), die sich mit ethnischen, geschlechtsspezifischen und altersspezifischen Unterschieden im Zusammenhang mit verschiedenen Variablen wie zum Beispiel *Ängstlichkeit*, *Erwartungshaltung*, *negativen Gedanken* beschäftigen. Im Rahmen dieser Studien wird *Stereotype Threat* jedoch nicht gemeinsam mit der *Theory of Mind* erfasst. In dieser Arbeit soll ein möglicher Zusammenhang zwischen der Bedrohung durch Stereotype und dem Bewältigen der *ToM*-Aufgaben hergestellt werden. Wie bereits angeführt handelt es sich bei der *ToM* um die Fähigkeit sich in andere hineinzusetzen und sowohl die eigenen Gefühle, Gedanken, Intentionen und Handlungen als auch die der anderen nachvollziehen und vorhersagen zu können. Die soziale Kompetenz, das soziale Verständnis und die Empathiefähigkeit sind einige der Kernkompetenzen der *ToM* und, wie bereits in mehreren Studien gezeigt werden konnte, bei Frauen stärker und besser ausgeprägt als bei Männern (Baron-Cohen & Wheelwright, 2004). Es könnte angenommen werden, dass das Erleben von *Stereotype Threat* mit dem Einfühlungsvermögen und der sozialen Kompetenz einer Person zusammenhängt, da die bedrohte Person die Fähigkeit besitzen muss, sich in die beurteilende Person hineinzusetzen, um deren Gedanken und Intentionen nachvollziehen zu können. Daraus könnte man schließen, dass Personen mit einer gut entwickelten *ToM* anfälliger auf die Bedrohung durch Stereotype und die dadurch verursachten Leistungseinbußen sind.

Ein weiteres Anliegen dieser empirischen Arbeit ist die Untersuchung des Konstrukts *Risikoverhalten*, im Zusammenhang mit *Stereotype Threat* und der *ToM*. Diesbezüglich gibt es Hinweise in der Studie von Carr und Steele (2009), dass *Stereotype Threat* risikoaversives- und vermeidendes Verhalten begünstigt.

In Bezug auf geschlechtsspezifische Unterschiede gelten Männer als risikoreicher und -freudiger in ihrem Entscheidungsverhalten im Gegensatz zu Frauen (Ben-Shakhar & Sinai, 1991; Byrnes et al.1999). Hinsichtlich eines Zusammenhangs mit der *ToM* liegen keine Ergebnisse vor.

Abschließend ist zu sagen, dass die Betrachtung der berichteten Forschungsergebnisse mögliche Hinweise auf potenzielle Verbindungen der drei Konstrukte liefert, die bis dato noch nicht in dieser Form untersucht worden sind. Es ist das Anliegen der vorliegenden Arbeit neue Ergebnisse zu dieser interessanten und bedeutenden Thematik zu generieren.

Kapitel 5 **Zielsetzungen, Fragestellungen und Hypothesen**

Im folgenden Kapitel wird einerseits auf das allgemeine Hauptziel der Untersuchung und andererseits auf die spezifischen Zielsetzungen, Fragestellungen und Hypothesen, die im Zuge dieser empirischen Arbeit verfolgt werden, eingegangen.

5.1. Zielsetzungen und Fragestellungen

Das prinzipielle Hauptziel dieser Studie ist die Untersuchung von sozialen und biologischen Einflussfaktoren auf die Leistung in Testsituationen bei Jugendlichen (Burschen und Mädchen). Ein längerfristiges und zukunftsorientiertes Anliegen und Ziel dieser Studie ist es den EMS-Test im Speziellen und Auswahlverfahren im Allgemeinen fairer für beide Geschlechter zu gestalten und/oder eine geschlechtsspezifische Normierung zu erzielen.

Das Ziel dieser empirischen Arbeit ist die Untersuchung der Theory of Mind in Abhängigkeit des Phänomens Stereotype Threat, Risikoverhalten und Geschlecht im Rahmen der Bewältigung von EMS-ähnlichen Aufgaben. Es sollen dabei mögliche Zusammenhänge und Unterschiede der drei Konstrukte unter Berücksichtigung geschlechtsspezifischer Unterschiede bei Burschen und Mädchen beleuchtet werden.

Diese Diplomarbeit behandelt einen Aspekt im Rahmen dieser größeren Untersuchung, die von der Vizerektorin für Personalentwicklung und Frauenförderung der Medizinischen Universität Wien, Univ.-Prof. Dr. Karin Gutiérrez-Lobos, in Zusammenarbeit mit dem Vizerektor für Studium und Lehre, Ao. Univ.-Prof. Dr.med.univ. Rudolf Mallinger und Univ.-Prof. Dr. Otto Scheiner vom Institut für Pathophysiologie und Allergieforschung, Leiter des Forschungssupport der Medizinischen Universität Wien angeregt wurde und unter der Leitung von Ao. Univ.- Prof. Mag. Dr. Ulrike Willinger (Universitätsklinik für Neurologie, Leitung des Arbeitsbereichs „Klinische- und Kognitive

Neuropsychologie“) und Mitarbeit der Diplomandinnen Gabriele Hangl, Helene Lager, Anita Teufl, Silvia Hameseder und Katrin Anzirk durchgeführt wurde. An dieser Stelle ist nochmals anzumerken, dass der idente Datensatz die allgemeine Grundlage für alle Berechnungen der Kolleginnen und der vorliegenden Arbeit darstellt, jedoch unterschiedliche Schwerpunkte verfolgt werden. Folgende Fragestellungen werden von den Kolleginnen behandelt:

- „Stereotype Threat im Zusammenhang mit der Risikobereitschaft in Bezug auf die Selbstwirksamkeitserwartung“ (vgl. Teufl, 2011)
- „Stereotype Threat im Zusammenhang mit der Risikobereitschaft in Bezug auf die Leistungsmotivation“ (vgl. Hangl, 2011)
- „ Geschlechtsunterschiede im Zusammenhang mit Stereotype Threat und der Risikobereitschaft in Bezug auf die Merkfähigkeit“ (vgl. Lager, 2011)
- „ Geschlechtsunterschiede im Zusammenhang mit Stereotype Threat und der Risikobereitschaft in Bezug auf Schlussfolgendes Denken“ (vgl. Hameseder, 2011)

5.2 Hypothesen

In diesem Abschnitt werden die aus den Fragestellungen abgeleiteten und formulierten Hypothesen dargestellt.

5.2.1 Haupthypothese

Es können signifikante Unterschiede in Abhängigkeit vom Geschlecht, den Stereotype Threat-Bedingungen und den Risikoverhaltes-Bedingungen im Bewältigen *Theory of Mind*-Aufgaben angenommen werden.

5.2.2 Hypothesen zur Theory of Mind

Geschlecht

H₁(1): Es kann ein signifikanter Unterschied in der ToM in Abhängigkeit vom Geschlecht angenommen werden.

Risikoverhalten

H₁(2): Es kann ein signifikanter Unterschied in der ToM in Abhängigkeit von den drei Ausprägungen des Risikoverhaltens angenommen werden.

Stereotype Threat-Bedingungen

H₁ (3): Es kann ein signifikanter Unterschied in der ToM in Abhängigkeit von den vier Stereotype Threat-Bedingungen angenommen werden.

Geschlecht und Stereotype Threat-Bedingung

H₁ (4): Es kann eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und von den Stereotype Threat Bedingungen angenommen werden.

Geschlecht und Risikoverhalten

H₁ (5): Es kann eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und den Ausprägungen des Risikoverhaltens angenommen werden.

Stereotype Threat und Risikoverhalten

H₁ (6): Es kann eine signifikante Interaktion in der ToM in Abhängigkeit von den Stereotype Threat-Bedingungen und dem Risikoverhalten angenommen werden.

Geschlecht, Risikoverhalten und Stereotype Threat

H₁ (7): Es kann eine signifikante Interaktion höherer Ordnung in der ToM in Abhängigkeit vom Geschlecht, von den Ausprägungen des Risikoverhaltens und den Stereotype Threat-Bedingungen angenommen werden.

5.2.2.1 Zusammenhänge zwischen der ToM und den EMS-Untertests

In diesem Hypothesenblock werden die jeweiligen Zusammenhänge zwischen der ToM und den einzelnen EMS-ähnlichen Dimensionen (*Quantitative und formale Probleme, Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Diagramme und Tabellen, Figuren lernen, Fakten lernen*) geprüft.

Haupthypothese:

.H₁ (8): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im jeweiligen Untertest angenommen werden.

Hypothesen zu den einzelnen EMS-Untertests:

H₁ (9): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Diagramme und Tabellen* angenommen werden.

H₁ (10): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Schlauchfiguren* angenommen werden.

H₁ (11): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Medizinisch-naturwissenschaftliches Grundverständnis* angenommen werden.

H₁ (12): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Diagramme und Tabellen* angenommen werden.

H₁ (13): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Figuren lernen* angenommen werden.

H₁ (14): Es kann ein signifikanter Zusammenhang zwischen der ToM und dem Abschneiden im Untertest *Fakten lernen* angenommen werden.

5.2.2.2 Nebenhypothese zur ToM

Wunsch, Medizin zu studieren

H₁ (15): Es kann ein signifikanter Unterschied in der ToM in Abhängigkeit von dem Wunsch, Medizin zu studieren (ja/nein) angenommen werden.

Geschlecht und Wunsch, nach Medizinstudium

H₁ (16): Es kann eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und dem Wunsch, Medizin zu studieren angenommen werden.

5.2.4 Nebenhypothese zum Game of Dice Task (GDT)

H₁ (17): Es kann ein signifikanter Unterschied im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit von der Risikoinstruktion angenommen werden.

H₁ (18): Es kann ein signifikanter Unterschied im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit vom Geschlecht angenommen werden.

H₁ (19): Es kann eine signifikante Interaktion im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit vom Geschlecht und der Risikoinstruktion angenommen werden.

II. Empirischer Teil

Kapitel 6 Methode

In diesem Kapitel soll die ursprüngliche Untersuchungsplanung, die Beschreibung der intendierten Stichprobe sowie der Vorstudie vorgestellt werden. Des Weiteren folgt eine kurze Darstellung und Beschreibung der eingesetzten Erhebungsinstrumente.

6.1 Untersuchungsplan und Intendierte Stichprobe

Wie bereits im Kapitel 5 erwähnt wurde, handelt es sich bei der vorliegenden Arbeit um einen Teil einer größer angelegten Untersuchung, die von Univ. Prof. Dr. Karin Gutiérrez-Lobos, der Vizerektorin für Personalentwicklung und Frauenförderung der Medizinischen Universität Wien, in Zusammenarbeit mit dem Vizerektor für Studium und Lehre, Ao. Univ.-Prof. Dr.med.univ. Rudolf Mallinger und Univ.-Prof. Dr. Otto Scheiner vom Institut für Pathophysiologie und Allergieforschung, Leiter des Forschungssupport der Medizinischen Universität Wien angeregt wurde und unter der Leitung von Ao. Univ.- Prof. Mag. Dr. Ulrike Willinger (Universitätsklinik für Neurologie, Leitung des Arbeitsbereichs „Klinische- und Kognitive Neuropsychologie“) und Mitarbeit der Diplomandinnen Gabriele Hangl, Helene Lagger, Anita Teufl, Silvia Hameseder und Katrin Anzirk durchgeführt wurde. Im Rahmen der Untersuchung wird von jeder Kollegin ein bestimmter Teilaspekt schwerpunktmäßig analysiert.

Ursprüngliches Ziel der Untersuchung war die Durchführung eines multivariaten Designs zur Berücksichtigung von biologischen Faktoren, wie dem Hormonstatus und sozialen Faktoren wie dem Phänomen Stereotype Threat und der Risikobereitschaft zur Klärung von geschlechtsspezifischen Unterschieden in kognitiven Leistungen mittels EMS-ähnlichen Aufgaben und ToM-Aufgaben unter Berücksichtigung von beeinflussenden und modulierenden *mediating*

variables (Mediatorvariablen) wie der Ängstlichkeit, Selbstwirksamkeit und Leistungsmotivation.

Die Studie sollte als experimentelles, multivariates Design mit randomisierter Zuteilung der VersuchsteilnehmerInnen zu den Versuchsbedingungen (Stereotype Threat, Risikoverhalten) durchgeführt werden.

Die Untersuchung besteht aus insgesamt 12 Versuchsbedingungskombinationen, die sich aus 24 vollständigen Faktorstufenkombinationen des festen Faktors Geschlecht (2) und der beiden experimentellen Faktoren Stereotype Threat (4), Risikoverhalten (3) ergeben (siehe Tabelle 2.).

Tabelle 2: Untersuchungsdesign

Variablen	Versuchsbedingungen											
	A			B			C			D		
Hormon- messung												
Stereotype Threat												
Risk Behavior	1	2	3	1	2	3	1	2	3	1	2	3
Geschlecht	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂
EMS- ähnl. Test												
Theory of Mind												
Hormon- messung												

Ängstlich- keit, Selbst- konzept, Leistungs- motivation																					
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Im Rahmen des Untersuchungsdesigns werden vier verschiedene Stereotype Threat-Bedingungen und drei unterschiedliche Instruktionen zum Risikoverhalten vorgegeben.

Stereotype-Threat-Bedingungen:

1. Aktivierung des Stereotype Threat

Die VersuchsteilnehmerInnen wurden darauf hingewiesen, dass Frauen bessere Leistungen im Gedächtnis erbringen und Männer eine Überlegenheit in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen, im Verstehen von komplexen Texten und im naturwissenschaftlichen Verständnis aufweisen.

2. Neutrale Bedingung

Es wurde betont, dass es keine geschlechtsspezifischen Unterschiede gibt.

3. Aufklärung

In dieser Bedingung wurden die VersuchsteilnehmerInnen über das Phänomen Stereotype Threat aufgeklärt.

4. Negativ-Aktivierung des Stereotype Threat

Es wurde, im Gegensatz zu der Aktivierungsbedingung, darauf hingewiesen, dass Frauen bessere Leistungen in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen, im Verstehen von komplexen Texten und im naturwissenschaftlichen Verständnis aufweisen und Männer bessere Leistungen im Gedächtnis erbringen.

Untersuchungsbedingungen für das Risikoverhalten:

1. *Instruktion: Raten*

Die UntersuchungsteilnehmerInnen wurden dahingehend instruiert, zu raten, sofern sie die richtige Lösung nicht wissen.

2. *Keine Instruktion:*

Es wurde angewiesen, die Aufgaben aufmerksam durchzulesen und zu bearbeiten.

3. *Instruktion:*

Den VersuchsteilnehmerInnen wurde mitgeteilt, dass sie eine Antwort nur dann ankreuzen sollten, wenn sie sich dessen ganz sicher sind.

Darüber hinaus wurden den UntersuchungsteilnehmerInnen EMS-ähnliche Aufgaben zu sechs der zehn EMS-Untertests mit begrenzter Bearbeitungszeit vorgegeben. Im Anschluss daran wurden die Testteilnehmer dazu aufgefordert den restlichen Teil, bestehend aus den ToM-Aufgaben, dem Leistungsmotivationsinventar und dem Selbstwirksamkeitserwartungsfragebogen, ohne zeitlicher Begrenzung fertig zu stellen.

Die Erhebung des Hormonstatus wurde zu zwei Zeitpunkten, vor und nach der Testung, mittels der EMS-ähnlichen Aufgaben, durch die Entnahme einer Speichelprobe durchgeführt. Darüber hinaus wurde das Datum der letzten Menstruation sowie Angaben bezüglich der Pilleneinnahme (ja/nein) erhoben.

Die Testung soll in österreichischen, bevorzugt in Wiener Gymnasien durchgeführt werden, für die die Autorinnen im Vorfeld Genehmigungen der jeweiligen Stadtschulräte, sowie das Einverständnis und die Zusage der Direktionen einholen mussten. Den VersuchsteilnehmerInnen wurde vor Beginn der Testung eine Einverständniserklärung ausgehändigt (siehe Anhang). Im Anschluss daran wurden die SchülerInnen gebeten, einen soziodemographischen Fragebogen (siehe Anhang) auszufüllen, in dem spezielle Interessenschwerpunkte, wie Ausbildung (AHS-Zweig), Ausbildung und Beruf

der Eltern, Wunsch nach einem Medizinstudium und die Beschäftigung mit EMS-Vorbereitung erfasst werden.

Insgesamt sollte die Testung zwei Unterrichtsstunden umfassen und in Gruppentestungen in den Klassenräumen der jeweiligen Schulen durchgeführt werden.

Der geplante Stichprobenumfang sollte aus N=1200 SchülerInnen (50 pro Versuchsbedingung) im Alter von 17-20 Jahren in 8. Klassen von Gymnasien (12.Schulstufe) bestehen, wobei diese schließlich über verschiedene Standorte und Ausbildungszweige (u.a. naturwissenschaftlich, humanistisch, neusprachlich) in Wien rekrutiert wurden. Geplanter Beginn der Untersuchung war ursprünglich Oktober 2009 und das voraussichtliche Ende war für Juni 2010 angesetzt.

6.2. Vorstudie

Im Zuge der Untersuchungsplanung wurde beschlossen, eine Vorstudie durchzuführen, um die Realisierbarkeit der Durchführung zu prüfen. Die Vorstudie wurde an zwei Terminen mit insgesamt N=32 MedizinstudentInnen des zweiten Abschnitts der Medizinischen Universität Wien durchgeführt. Das Ziel der Vorstudie war abzuklären und zu überprüfen, ob und inwieweit die experimentelle Variation der vier Stereotype Threat-Bedingungen sowie der drei Untersuchungsbedingungen zum Risikoverhalten greift und sich daraus Auswirkungen oder Effekte daraus ableiten lassen. Darüber hinaus sollten potenzielle Deckeneffekte abgeklärt werden, da die MedizinstudentInnen den EMS bereits absolviert hatten. Des Weiteren sollte die zeitliche und praktische Anwendbarkeit der Aufgabenstellungen in dem von den Autorinnen zusammengestellten Testheft überprüft werden.

Im Zuge der Durchführung der Vorstudie ergaben sich folgende Änderungen für die Untersuchungsdurchführung:

- Streichung des Untertests *Textverständnis*
- Erweiterung des Wissenstests mit Selbsteinschätzungsaufgaben

6.3. Erhebungsinstrumente

Im folgenden Abschnitt werden die in der Untersuchung eingesetzten Inventare dargestellt.

6.3.1 Soziodemografischer Fragebogen

Der soziodemografische Fragebogen wurde zur Erhebung der interessierenden und für die Untersuchung relevanten Daten (wie z.B. geplante Studienwahl, Berufswunsch, Beschäftigung mit dem EMS etc.) konstruiert. Er besteht aus 10 Fragen, die sowohl ein offenes als auch ein geschlossenes zwei-bis-achtkategorielles Antwortformat aufweisen.

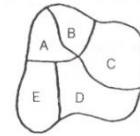
6.3.2 Wissenstest

Der *Wissenstest* wurde, in Anlehnung an das *Beliefs About the Cognitive Abilities of Females and Males Questionnaire* von Halpern und Tan (2001), zur Überprüfung der Aktivierungswirksamkeit der vier Stereotype Threat-Bedingungen erstellt und setzt sich aus insgesamt 14 Aufgaben zusammen, wobei jeweils sieben Fremdeinschätzungs- und sieben Selbsteinschätzungsfragen zu beantworten sind. Im Rahmen der sieben Fremdeinschätzungsaufgaben werden die Testpersonen instruiert ihre Einschätzung, ob es sich im vorgegebenen Beispiel um eine Frau oder um einen Mann handelt, in Form von Prozentangaben einzutragen, sodass die Summe 100 Prozent ergibt. Zur Veranschaulichung werden im Anschluss Beispielimbeispiele dargestellt (siehe Abbildungen 1 - 4).

Beispielitem 1:

Denken Sie an eine Person, die Sie nicht kennen. Die folgende Aufgabe, welche prüft wie gut Einzelheiten von Gegenständen wahrgenommen, eingepägt und gemerkt werden können wurde richtig gelöst.

Einprägen der schwarzen Flächen verschiedener Figuren:



Welche Fläche war die geschwärzte?

richtige

Lösung: (C)

Diese Aufgabe wurde richtig gelöst, die Wahrscheinlichkeiten, **dass es sich um einen Mann bzw. Frau handelt, betragen:**

Mann _____% bzw. Frau _____ % (Summenwert soll 100% ergeben)

Abbildung 1: Beispielitem 1- Fremdeinschätzung

Beispielitem 2: Fremdeinschätzungsaufgabe zur ToM

7) Denken Sie an eine Person, die Sie nicht kennen. Diese Person kann sich sehr gut in andere Menschen hineinversetzen und deren Gedanken und Probleme nachvollziehen.

Die Wahrscheinlichkeiten, **dass es sich um einen Mann bzw. Frau handelt, betragen:**

Mann _____% bzw. Frau _____ % (Summenwert soll 100% ergeben)

Abbildung 2: Beispielitem 2- Fremdeinschätzung ToM

Im Anschluss an die Fremdeinschätzungsaufgaben werden den UntersuchungsteilnehmerInnen sieben Selbsteinschätzungsaufgaben vorgegeben,

bei denen sich die Testpersonen in Bezug auf eine genannte Fähigkeit einschätzen müssen.

Beispielitem 3: Selbsteinschätzungsaufgabe

Denken Sie an folgende Aufgabe: Einer Person wurde ein Text zu einem medizinisch-naturwissenschaftlichen Thema (z.B. Herzkreislaufsystem) vorgelegt zu dem sie anschließend Verständnisfragen beantworten musste.						
Ich verfüge über die Fähigkeit, die oben angeführte Aufgabe richtig zu lösen:						
1	2	3	4	5	6	7
trifft nicht zu					trifft sehr zu	

Abbildung 3: Beispielitem 3 – Selbsteinschätzung

Beispielitem 4: Selbsteinschätzung zur ToM

Denken Sie an eine Person, die sie nicht kennen. Diese Person kann sich sehr gut in andere Menschen hineinversetzen und deren Gedanken und Probleme nachvollziehen.						
Ich verfüge über die Fähigkeit, die oben angeführte Aufgabe richtig zu lösen:						
1	2	3	4	5	6	7
trifft nicht zu					trifft sehr zu	

Abbildung 4: Selbsteinschätzung ToM

6.3.3 EMS-Untertests

Der Eignungstest für das Medizinstudium (EMS) besteht aus 10 Untertests, die jeweils zu einem Gesamtwert verrechnet werden. Im Rahmen dieser Untersuchung wurden jedoch nicht alle 10 Untertests, sondern nur eine Auswahl von sechs Untertests mit vierstufigem Antwortformat vorgegeben. Im folgenden Abschnitt werden die einzelnen Untertests vorgestellt, in dem einerseits Angaben zum ursprünglichen Aufgaben- und Zeitumfang sowie zu den Reliabilitäten der

Untertests des EMS und andererseits Angaben bezüglich der für diese Untersuchung relevanten Veränderungen und Kürzungen der jeweiligen Untertests angeführt werden.

In der folgenden Tabelle 3 werden die Reliabilitäten nach Cronbachs Alpha für die zehn EMS-Untertests und deren Entwicklung von 1998-2007 dargestellt (Hänsgen & Spicher, 2007).

Tabelle 3: Reliabilitäten nach Cronbachs Alpha (α) für die EMS-Untertests von 1998-2007 nach Hänsgen und Spicher (2007).

Untertest	Innere Konsistenz (Cronbachs Alpha)										
	Deutschland	Schweiz (d)									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Muster zuordnen	.64 - .69	.69	.60	.53	.56	.57	.53	.54	.67	.60	.55
Med.-nat. Grundverständnis	.64 - .73	.70	.70	.61	.66	.57	.72	.66	.72	.71	.65
Schlauchfiguren	.76 - .80	.71	.74	.71	.73	.68	.74	.72	.72	.70	.73
Quant. u. form. Probleme	.69 - .71	.76	.75	.76	.74	.71	.67	.71	.76	.71	.70
Textverständnis	.68 - .71	.73	.75	.74	.70	.68	.78	.67	.73	.66	.69
Figuren lernen	.69 - .74	.72	.70	.62	.64	.82	.77	.71	.74	.66	.63
Fakten lernen	.62 - .70	.68	.72	.64	.69	.70	.73	.72	.72	.73	.73
Diagramme und Tabellen	.67 - .71	.67	.75	.63	.61	.62	.70	.66	.70	.61	.69
Planen und Org.									.62	.42	.63

Tabelle 45: Zuverlässigkeit 1998 bis 2007; Vergleich der Skalenkonsistenzen für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

In der Tabelle 4 werden nochmals gesondert die Reliabilitäten nach Cronbachs Alpha (α) für die in der Testung eingesetzten sechs EMS-Untertests von Hänsgen und Spicher (2007) dargestellt.

Tabelle 4: Reliabilitäten der sechs eingesetzten EMS-Untertests (Hänsgen & Spicher, 2007)

EMS-Untertests	Reliabilität (Cronbachs Alpha α)
Quantitative und formale Probleme	.70
Schlauchfiguren	.73
Medizin.- nat.-wissenschaftl. GV.	.65

Diagramme und Tabellen	.69
Figuren lernen	.63
Fakten lernen	.73

6.3.3.1 Quantitative und formale Probleme

Der Untertest *Quantitative und formale Probleme* erfasst den Umgang mit Zahlen und Formeln im Rahmen medizinisch-naturwissenschaftlicher Fragen und besteht aus 20 Aufgaben für die eine Bearbeitungszeit von 50 min zur Verfügung steht.

Tatsächliche Vorgabe der EMS-ähnlichen Aufgaben: 16 Aufgaben mit einer Bearbeitungszeit von 10 min.

6.3.3.2 Schlauchfiguren

Hierbei handelt es sich um Test zur Erfassung des räumlichen Vorstellungsvermögens bei Jugendlichen und jungen Erwachsenen. Dieser Untertest besteht aus 21 Aufgaben, die mit einer begrenzten Bearbeitungsdauer von 12min vorgegeben werden.

Tatsächliche Vorgabe: 21 Aufgaben mit einer Bearbeitungszeit von 12 min.

6.3.3.3 Medizinisch-naturwissenschaftliches Grundverständnis

Der Untertest *Medizinisch-naturwissenschaftliches Grundverständnis* erfasst das Verständnis für Fragen der Medizin und der Naturwissenschaften und umfasst 20 Aufgaben für die eine Bearbeitungszeit von 50 min vorgesehen ist.

Tatsächliche Vorgabe: 16 Aufgaben mit einer Bearbeitungszeit von 10 min.

6.3.3.4 Diagramme und Tabellen

Dieser Untertest erfasst die Analyse- und Interpretationsfähigkeit von Diagrammen und Tabellen und umfasst ebenfalls 20 Aufgaben, die in einer Bearbeitungszeit von 50 min absolviert werden sollen.

Tatsächliche Vorgabe: 16 Aufgaben mit einer vorgesehenen Bearbeitungszeit von 10 min.

6.3.3.5 Figuren und Fakten lernen

Diese beiden Untertests werden gemeinsam, zuerst in einer Einprägephase von jeweils 4 und 6 min und dann in einer Reproduktionsphase von jeweils 5 und 7 min im Umfang von 20 Aufgaben pro Untertest vorgegeben. Der Untertest *Figuren lernen* erfasst das Wahrnehmen und Einprägen von visuell dargebotenen Einzelheiten. Der Untertest *Fakten lernen* erfasst das Einprägen von Fakten.

Tatsächliche Vorgabe: Figuren lernen: Einprägephase 2 Minuten

Reproduktion 2 Minuten

Tatsächliche Vorgabe: Fakten lernen Einprägephase: 3 Minuten

Reproduktionsphase: 3 Minuten

6.3.4 Leistungsmotivationsinventar (LMI)

Das Leistungsmotivationsinventar von Schuler et al. (2001) ist ein Einstellungsfragebogen mit einem siebenstufigen Antwortformat zur Erfassung der beruflichen Leistungsmotivation, bestehend aus 170 Items, die 17 Dimensionen zugeordnet sind. Eckardt und Schuler (1992) stufen die allgemeine Leistungsmotivation, neben dem Vorhandensein der grundlegenden kognitiven Fähigkeiten als zweitwichtigstes und für den beruflichen Erfolg ausschlaggebendes Merkmal ein.

Im Rahmen dieser Untersuchung werden zwei ausgewählte Skalen des LMI eingesetzt, die Internalität und Wettbewerbsorientierung erfassen.

Internalität: "Gemäß dem auf Rotter (1966) zurückgehenden Konzept locus of control (dt. Kontrollüberzeugung, s. Krampen, 1991) wird unter Internalität die generalisierte Überzeugung verstanden, dass Handlungsergebnisse tendenziell eher als selbstverursacht und selbstverantwortet erlebt werden, anstatt sie externalen Ursachen zuzuschreiben. Nach den Skalenanalysen fällt Internalität mit der zunächst hiervon unterschiedenen Attributionsneigung sensu Weiner und Kukla (1970) zusammen. Personen mit hohen Werten sind dadurch gekennzeichnet, dass sie Ergebnisse und Konsequenzen ihres Verhaltens auf internale Ursachen zurückführen, Sie sind der Überzeugung, dass das meiste im Leben von ihnen selbst, ihrem eigenen Verhalten und Kräfteinsatz abhängt. Beruflicher Erfolg ist für sie keine Glückssache, für Misserfolge und Minderleistung fühlen sie sich selbst verantwortlich."(S.15)

Wettbewerbsorientierung: "Wettbewerbsorientierung erfasst die Tendenz, Konkurrenz als Ansporn und Motivation für berufliche Leistung zu erleben. Personen mit hohen Werten suchen den Wettbewerb und den Vergleich mit anderen. Sie wollen besser und schneller sein als andere. Gewinnen ist ihnen wichtig und bestärkt sie in ihrer Anstrengung. Wettbewerbsorientierung als Konzept stammt aus unterschiedlichen Ansätzen zur Leistungsmotivationstheorie, verschiedentlich allerdings bezogen auf die Auseinandersetzung mit sich selbst, nicht mit Konkurrenten (z.B. McClelland, 1985; Cassidy & Lynn, 1989)." (S.16)

6.3.5 Skala zur Allgemeinen Selbstwirksamkeitserwartung

Zur Erfassung der Selbstwirksamkeitserwartung wurde die *Skala zur Allgemeinen Selbstwirksamkeitserwartung* von Schwarzer und Jerusalem (1999) verwendet. Die Skala besteht aus 10 Items mit einem vierstufigen Antwortformat.

„Das Konzept der Allgemeinen Selbstwirksamkeitserwartung fragt nach der persönlichen Einschätzung der eigenen Kompetenzen, *allgemein* mit Schwierigkeiten und Barrieren im täglichen Leben zurechtzukommen.

Selbstwirksamkeitserwartungen beeinflussen vor allem die *Auswahl* von Handlungen (Schwierigkeitsgrad), die investierte *Anstrengung* im Zielerreichungsprozess, die *Ausdauer* angesichts von Schwierigkeiten und Barrieren sowie indirekt den Grad den *Handlungserfolgs*.“ (Schwarzer & Jerusalem, 1999, S. 57).

6.3.6 ToM-Aufgaben

Bei den *ToM-Aufgaben* handelt es sich um einen Test zur Erfassung der ToM im Sinne einer *False Belief erster, zweiter und dritter Ordnung* und wird in Form von drei konstruierten Geschichten vorgelegt (Willinger et al., in Vorbereitung). Den VersuchsteilnehmerInnen werden im Rahmen dieser drei Geschichten spezifische Situationen dargestellt zu denen sie im Anschluss Fragen beantworten müssen. Das Antwortformat der *ToM-Aufgaben* ist zwei- drei- und vierkategorial (siehe Abbildung 5,6,7), wobei es nur eine richtige Antwortalternative gibt. Die gelösten Items fließen ungewichtet in den ToM-Gesamtscore ein. Zur Veranschaulichung werden infolge Beispiele für jede der drei Geschichten angeführt (siehe Abbildung 5,6,7).

Kurze Definition der *Theory of Mind*:

Es handelt sich dabei um „die grundlegende Fähigkeit sich in andere hineinzusetzen, ihre Gefühle, ihr Denken und Wissen einzubeziehen und zu erkennen, dass die Handlungen anderer Menschen durch ihre mentalen und emotionalen Zustände bedingt sind.“ (Ferstl , 2007, S.68) Des Weiteren versteht man darunter alltagspsychologische Konzepte, die man einsetzt, um sich selbst und den anderen mentale Zustände zuzuschreiben (Sodian, 2007).

False Beliefs erster, zweiter und dritter Ordnung

False Belief erster Ordnung: Es handelt sich dabei um Aufgaben zu Repräsentation einer Überzeugung über einen Zustand der Welt. Im Alter von vier Jahren können Kinder Aufgaben der ersten Ordnung lösen und haben ein Verständnis dafür entwickelt, dass Menschen Überzeugungen haben, von denen das Kind weiß, dass sie falsch sind.

False Belief zweiter Ordnung: Zwischen vier und sechs Jahren finden erhebliche Erweiterungen und Differenzierungen statt. In diesem Alter entwickeln Kinder das Verständnis für Ironie und Witz und den Gebrauch von Metaphern. Es handelt sich dabei um Aufgaben, die die Überzeugung einer Person A über die Überzeugung einer Person B erfordern. Das Verständnis, dass eine Überzeugung über eine Überzeugung einer anderen Person falsch sein kann.

False Belief dritter Ordnung: Hierbei handelt es sich um Aufgaben über die Absicht über eine falsche Überzeugung über eine Absicht.

Beispiele:

Erste Geschichte: Im Büro

Elsa Prohaska und Franz Schneider sitzen im Büro und bereiten diverse Unterlagen für eine Besprechung mit einer wichtigen Kundin vor. Als sie damit fertig sind legen sie die Unterlagen auf Hrn. Schneiders Schreibtisch. Bevor beide das gemeinsame Arbeitszimmer verlassen, um andere dringende Aufgaben zu erledigen, vereinbaren sie, dass Fr. Prohaska kurz vor der Besprechung die Unterlagen holt und diese in den Konferenzraum mitnimmt, wo die Besprechung stattfinden wird.

In der Zwischenzeit sieht Anna Meier, eine weitere Mitarbeiterin die Besprechungsunterlagen und will sie Hrn. Schneider bringen, der sich schon im Konferenzraum befindet, um die dortige technische Ausstattung für die Besprechung zu überprüfen. Gleichzeitig macht sich Fr. Prohaska auf den Weg, die Unterlagen zu holen.

Frage: Wo wird Fr. Prohaska die Unterlagen suchen?

im Arbeitszimmer im Konferenzraum

Abbildung 5: Im Büro-zweikategoriales Antwortformat

Zweite Geschichte: Briefverkehr

Auf dem Weg zu den diktierten Kassetten trifft Hr. Mitterer Gernot, einen langjährigen Kollegen, der ihm mitteilt, dass Fr. Cerwinka die Kassetten gerade ins Schreibbüro getragen hat.

Frage: Weiß Fr. Cerwinka, dass Hr. Mitterer weiß, wo die diktierten Kassetten sind?

ja nein

Frage: Wo denkt Fr. Cerwinka, dass Hr. Mitterer die diktierten Kassetten suchen wird?

im Chefzimmer im Sekretariat im Schreibbüro

Abbildung 6: Briefverkehr – dreikategorielles Antwortformat

Dritte Geschichte: In der Werbeagentur

Auf dem Weg zum Werbeplakat trifft Hr. Gerstl Jonathan, den Ferialpraktikanten, der ihm erzählt, dass er auf den Weg in den Kreativraum-2 ist, um Hr. Krämmer beim Ausfertigen des Werbeplakats zu unterstützen.

Frage: Weiß Hr. Krämmer, dass Hr. Gerstl weiß, wo das Werbeplakat ist?

ja nein

Frage: Wo denkt Hr. Krämmer, dass Hr. Gerstl das Werbeplakat suchen wird?

im Zimmer-Gerstl in Kreativwerkstätte-1 im Zimmer-Krämmer in Kreativwerkstätte-2

Abbildung 7: In der Werbeagentur- vierkategorielles Antwortformat

(Anmerkung: Bei Interesse für die ToM-Aufgaben können diese bei Univ.- Prof. Mag. Dr. Ulrike Willinger angefordert werden.)

6.3.7 Game of Dice Task (GDT)

Der *Game of Dice Task* (GDT, Brand, Fujiwara et al. 2005) ist ein objektives Verfahren, in Form eines computerisierten Würfelspiels, zur Erfassung des Risikoverhaltens und der Risikofreude. Die VersuchsteilnehmerInnen werden angewiesen, in 18 Durchgängen, mit fiktivem Geld (Startguthaben 1000 €) auf das Ergebnis eines Würfelwurfs zu wetten, in dem sie eine von mehreren Alternativen wählen, die sich hinsichtlich der Gewinnwahrscheinlichkeit und der Gewinn- bzw. Verlusthöhe unterscheiden. Die Testpersonen haben die Möglichkeit entweder auf eine, zwei, drei oder vier Augenzahlen zu setzen, wonach sich ihre Gewinnchance erhöht oder reduziert.

Kapitel 7 Untersuchung:

Im folgenden Kapitel wird auf die Untersuchungsdurchführung und die angewendeten Auswertungsverfahren eingegangen. Im Anschluss daran erfolgt die deskriptive Beschreibung der Gesamtstichprobe.

7.1. Untersuchungsdurchführung

Ursprünglich sollte, wie bereits im Abschnitt 6.1 der Untersuchungsplanung erwähnt, die Untersuchung im Oktober 2009 starten. Es kam jedoch zu einigen Verzögerungen, da einerseits der positive Bescheid der Ethikkommission abgewartet werden musste und andererseits die Genehmigungen der unterschiedlichen Stadtschulräte einzuholen waren. Des Weiteren erwies es sich als relativ zeitintensiv, Zusagen und Testungstermine von den Direktionen der einzelnen Schulen zu erhalten. Aufgrund dieser Verzögerungen konnten die Autorinnen letztendlich im Jänner 2010 mit der Untersuchung starten. Die letzte Testung fand Anfang Februar 2011 statt. Die Testungen wurden in den Klassenräumen der jeweiligen Schulen durchgeführt und umfassten zwei Unterrichtsstunden. Vor Beginn der Testung wurde den SchülerInnen eine Einverständniserklärung ausgehändigt und sie wurden in groben Zügen über den zeitlichen Ablauf der Testung informiert. Da im Rahmen der Testung die Zeit für die jeweiligen Untertests gestoppt wurde, gab es zwischen den Unterrichtsstunden keine Pause. Die Hormonmessung konnte aus Gründen beschränkter finanzieller und materieller Ressourcen nicht an allen Schulen durchgeführt werden. Die Entnahme der Speichelprobe für die Erhebung des Hormonstatus wurde jeweils vor und nach der Testung durchgeführt. Im Anschluss an die Testung wurde den SchülerInnen ein fünfminütiger Computertest, der Game of Dice Task (GDT), zur Erfassung der Risikobereitschaft, vorgegeben.

7.2 Stichprobe

In diesem Abschnitt wird die untersuchte Stichprobe in Bezug auf interessierende Aspekte wie die Geschlechter- und Altersverteilung, besuchte Schulzweige und die geplante Studienwahl beschrieben. Im Anschluss daran erfolgt eine dekodiert-statistische Darstellung der Erhebungsinstrumente.

7.2.1 Geschlechterverteilung innerhalb der Gesamtstichprobe

Insgesamt konnten 1093 SchülerInnen der 12. Schulstufe in Gymnasien in Wien (72,7 %), Burgenland (7,3 %) und Niederösterreich (20,0 %) untersucht werden. Davon waren 409 (37,4 %) der Untersuchten männlich und 684 (62,6 %) Testpersonen weiblich (siehe Tabelle 5)

Tabelle 5: Häufigkeiten in Abhängigkeit von Geschlecht

Geschlecht		Häufigkeit	Prozent	Gültige Prozent
Gültig	männlich	409	37,4	37,4
	weiblich	684	62,6	62,6
	Gesamt	1093	100,0	

In Abbildung 8 wird die Geschlechterverteilung der Gesamtstichprobe nochmals in Prozentwerten dargestellt.

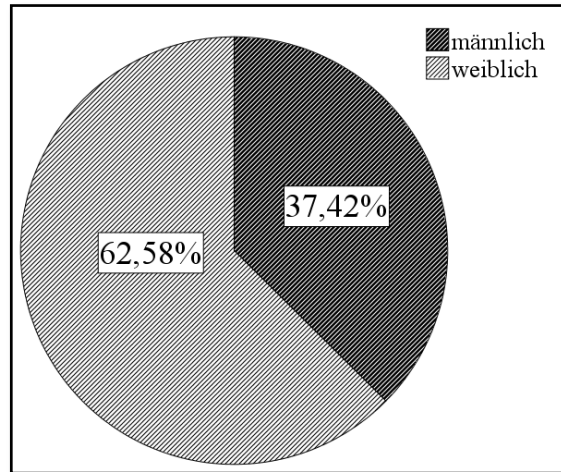


Abbildung 8: Prozentwerte der Geschlechter

In Tabelle 6 wird die Signifikanzprüfung der Geschlechterverteilung dargestellt.

Tabelle 6: Signifikanzbeurteilung der Geschlechterverteilung

Chi ² -Test	
	Geschlecht
$\chi^2 (1)$	69,190 ^a
Asymptotische Signifikanz	$p \leq .0001$

Es kann ein signifikanter Verteilungsunterschied im Geschlecht angenommen werden. Die Stichprobe weist überzufällig mehr Mädchen als Burschen auf.

7.2.1.1 Geschlechterverteilung in allen Versuchsbedingungen

Im Folgenden wird die Verteilung der UntersuchungsteilnehmerInnen auf die Versuchsbedingungen (Stereotype Threat, Risikoverhalten) dargestellt. Eingangs werden jeweils die Verteilungen der gesamten SchülerInnen auf die Versuchsbedingungen angeführt.

Stereotype Threat

Aus Tabelle 7 sind die Häufigkeiten ablesbar, mit denen die UntersuchungsteilnehmerInnen auf die vier Stereotype Threat-Bedingungen

aufgeteilt sind. Es zeigt sich bereits eine annähernde Gleichverteilung der SchülerInnen auf die vier Bedingungen.

Tabelle 7: Verteilung der SchülerInnen auf die vier Stereotype Threat-Bedingungen

Stereotype Threat			
		Häufigkeit	Prozent
Gültig	A	293	26,8
	B	257	23,5
	C	280	25,6
	D	263	24,1
	Gesamt	1093	100,0

Legende: A = Aktivierung

B = Neutral

C = Aufklärung

D = Negativ-Aktivierung

Tabelle 8: Signifikanzprüfung der Verteilung auf die Stereotype Threat Bedingungen

Chi ² -Test	
	Stereotype
$\chi^2(3)$	2,945 ^a
Asymptotische Signifikanz	<i>p</i> = .400

Die Berechnung der entsprechenden Prüfgröße ergibt ein nicht signifikantes Ergebnis (siehe Tabelle 8). Es kann kein signifikanter Verteilungsunterschied in den Stereotype Threat-Bedingungen angenommen werden.

Aus der Kontingenztafel (Tabelle 9) kann die Verteilung der UntersuchungsteilnehmerInnen auf die vier Stereotype Threat-Bedingungen in Abhängigkeit vom Geschlecht entnommen werden.

Mittels Kontingenztafeln (Kreuztabellen) können Zusammenhänge zwischen zwei nominalskalierten Merkmalen geprüft werden. Es erfolgen ein Vergleich mit theoretischen Verteilungen, sowie die Überprüfung, ob sich die empirischen von den theoretischen Häufigkeiten unterscheiden. Bei einer gegebenen Assoziation zwischen zwei Variablen kann mit Hilfe des χ^2 -Test überprüft werden, ob ein signifikanter Zusammenhang besteht (Bortz & Döring, 2006, S.613).

Tabelle 9: Kontingenztafel Geschlecht und Stereotype Threat

Kreuztabelle – Geschlecht * Stereotype Threat							
			Stereotype				Gesamt
			A	B	C	D	
Geschlecht	männlich	Anzahl	123	85	112	89	409
		Erwartete Anzahl	109,6	96,2	104,8	98,4	409,0
		% innerhalb	30,1%	20,8%	27,4%	21,8%	100,0%
		Standardisierte Residuen	1,3	-1,1	,7	-,9	
	weiblich	Anzahl	170	172	168	174	684
		Erwartete Anzahl	183,4	160,8	175,2	164,6	684,0
		% innerhalb	24,9%	25,1%	24,6%	25,4%	100,0%
		Standardisierte Residuen	-1,0	,9	-,5	,7	
Gesamt		Anzahl	293	257	280	263	1093
		Erwartete Anzahl	293,0	257,0	280,0	263,0	1093,0
		% innerhalb	26,8%	23,5%	25,6%	24,1%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(3) = 6.909$, $p = .075$ ergibt kein signifikantes Ergebnis, woraus folgt, dass kein Verteilungsunterschied in den Stereotype Threat-Bedingungen in Abhängigkeit vom Geschlecht zu Grunde liegt. Es kann eine Gleichverteilung der SchülerInnen auf die Stereotype Threat-Bedingungen angenommen werden.

Risikoverhalten

Aus Tabelle 10 ist ersichtlich wie viele SchülerInnen sich in der jeweiligen Risikobedingung befinden. Auch hier kann man bereits anmerken, dass eine annähernde Gleichverteilung der UntersuchungsteilnehmerInnen auf die drei Risikobedingungen angenommen werden kann.

Tabelle 10: Verteilung der SchülerInnen auf die drei Risikobedingungen

Risikoverhalten			
		Häufigkeit	Prozent
Gültig	raten	373	34,1
	ohne Instruktion	368	33,7
	sicher	352	32,2
	Gesamt	1093	100,0

In Tabelle 11 ist die Signifikanzprüfung der Verteilung der UntersuchungsteilnehmerInnen in den Risikobedingungen dargestellt.

Tabelle 11: Signifikanzprüfung der Geschlechterverteilung in den Risikobedingungen

Chi ² Test	
	Risikoverhalten
$\chi^2 (2)$,661 ^a
Asymptotische Signifikanz	$p=.719$

Die Signifikanzprüfung ergibt ein nicht signifikantes Ergebnis. Es kann somit kein signifikanter Verteilungsunterschied
In der folgenden Kontingenztafel (Tabelle 12) zeigt die Verteilung der SchülerInnen auf die drei Risikobedingungen in Abhängigkeit vom Geschlecht.

Tabelle 12: Kontingenztafel Geschlecht und Risikoverhalten

Kreuztabelle – Geschlecht * Risikoverhalten						
			Risikoverhalten			Gesamt
			raten	ohne Instruktion	sicher	
Geschlecht	männlich	Anzahl	145	131	133	409
		Erwartete Anzahl	139,6	137,7	131,7	409,0
		% innerhalb von Geschlecht	35,5%	32,0%	32,5%	100,0%
		Standardisierte Residuen	,5	-,6	,1	
	weiblich	Anzahl	228	237	219	684
		Erwartete Anzahl	233,4	230,3	220,3	684,0
		% innerhalb von Geschlecht	33,3%	34,6%	32,0%	100,0%
		Standardisierte Residuen	-,4	,4	-,1	
Gesamt		Anzahl	373	368	352	1093
		Erwartete Anzahl	373,0	368,0	352,0	1093,0
		% innerhalb von Geschlecht	34,1%	33,7%	32,2%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(2) = .878$, $p = .645$ ergibt kein signifikantes Ergebnis. Es können keine signifikanten Verteilungsunterschiede in den Risikobedingungen in Abhängigkeit vom Geschlecht angenommen werden.

Stereotype Threat und Risikoverhalten

In der nachfolgenden Tabelle 13 werden die Verteilungen der TestteilnehmerInnen auf die Risikobedingungen in Abhängigkeit von den Stereotype Threat-Bedingungen dargestellt.

Tabelle 13: Kontingenztafel Stereotype Threat und Risikoverhalten

Kreuztabelle – Stereotype Threat * Risikoverhalten						
			Risikoverhalten			Gesamt
			raten	ohne Instruktion	sicher	
Stereotype	A	Anzahl	93	108	92	293
		Erwartete Anzahl	100,0	98,6	94,4	293,0
		% innerhalb von Stereotype	31,7%	36,9%	31,4%	100,0%

		Standardisierte Residuen	-,7	,9	-,2	
B	Anzahl	90	87	80	257	
	Erwartete Anzahl	87,7	86,5	82,8	257,0	
	% innerhalb von Stereotype	35,0%	33,9%	31,1%	100,0%	
	Standardisierte Residuen	,2	,1	-,3		
C	Anzahl	104	92	84	280	
	Erwartete Anzahl	95,6	94,3	90,2	280,0	
	% innerhalb von Stereotype	37,1%	32,9%	30,0%	100,0%	
	Standardisierte Residuen	,9	-,2	-,7		
D	Anzahl	86	81	96	263	
	Erwartete Anzahl	89,8	88,5	84,7	263,0	
	% innerhalb von Stereotype	32,7%	30,8%	36,5%	100,0%	
	Standardisierte Residuen	-,4	-,8	1,2		
Gesamt	Anzahl	373	368	352	1093	
	Erwartete Anzahl	373,0	368,0	352,0	1093,0	
	% innerhalb von Stereotype	34,1%	33,7%	32,2%	100,0%	

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 5.121$, $p = .528$ zeigt ein nicht signifikantes Ergebnis. Es können keine Verteilungsunterschiede im Risikoverhalten in Abhängigkeit von Stereotype Threat angenommen werden.

Abschließend kann festgehalten werden, dass die Verteilung der SchülerInnen auf einzelne Faktorstufenkombinationen als nicht überzufällig häufig bezeichnet werden kann.

7.2.2 Alter der SchülerInnen

Zum Erhebungszeitpunkt waren die SchülerInnen ($N = 1086$ gültige Fälle) zwischen 16 und 21 Jahren alt ($M = 17$, $SD = 0.76$), wobei die 17-jährigen mit 580 Untersuchten (53,4%) anteilmäßig die größte Gruppe darstellen. Die Gruppe der 21-jährigen ist mit vier TeilnehmerInnen (0,37%) am geringsten bei den SchülerInnen vertreten (siehe Abbildung 9). Von den 1093 getesteten TestteilnehmerInnen, enthielten sich sieben Personen einer Altersangabe.

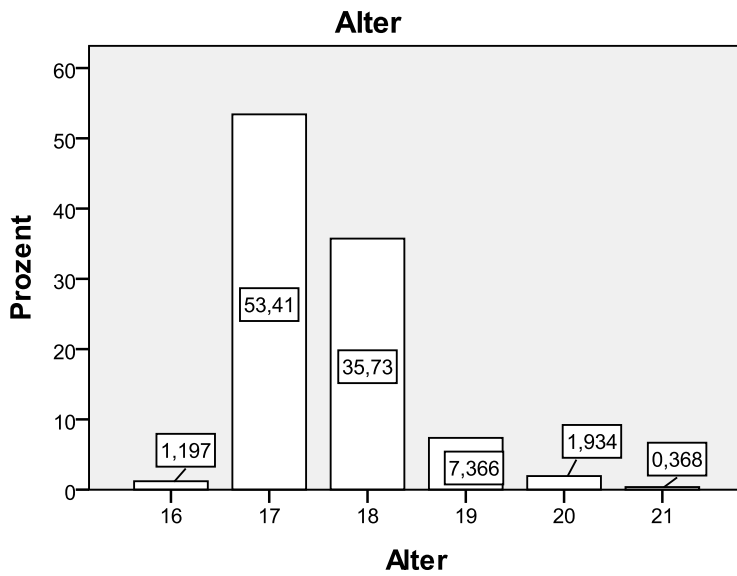


Abbildung 9: Häufigkeitsverteilung (Anteilswerte in Prozent) der Altersstufen (1086 gültige Fälle)

In Tabelle 14 werden die entsprechenden Häufigkeiten dargestellt.

Tabelle 14: Häufigkeiten der Altersstufen

Alter			
	Beobachtetes N	Erwartete Anzahl	Residuum
16	13	181,0	-168,0
17	580	181,0	399,0
18	388	181,0	207,0
19	80	181,0	-101,0
20	21	181,0	-160,0
21	4	181,0	-177,0
Gesamt	1086		

Die Signifikanzprüfung ergibt mit $p \leq .0001$ ein signifikantes Ergebnis, was bedeutet, dass keine Gleichverteilung des Alters angenommen werden kann (siehe Tabelle 15).

Tabelle 15: Signifikanzprüfungsprüfung Alter

Chi²-Test: Alter	
	Alter
Chi-Quadrat	1643,116 ^a
df	5
Asymptotische Signifikanz	$p \leq .0001$

7.2.3 Schulzweige und Staatsbürgerschaft

Bezüglich der Staatsbürgerschaftszugehörigkeit liegen von 1050 SchülerInnen Angaben vor. Die anteilmäßig größte Gruppe der getesteten SchülerInnen stellen mit 992 (94,5 %) Personen die ÖsterreicherInnen dar. Die restlichen 58 (5,5%) TestteilnehmerInnen gaben eine ausländische Staatsbürgerschaft an, wie Deutschland, Serbien, Slowakei, Ungarn, Rumänien, Polen, Japan, Kroatien, Schweden und Türkei.

Aus Tabelle 16 ist ersichtlich, welche die in der Stichprobe am häufigsten besuchten Schulzweige der UntersuchungsteilnehmerInnen der Oberstufe darstellen. Es zeigt sich, dass der Großteil der 1087 TestteilnehmerInnen, von denen Angaben diesbezüglich vorliegen, mit 288 Personen (26,5 %) den sprachlichen Schulzweig besucht, gefolgt von 247 SchülerInnen (22,7%), die angeben den naturwissenschaftlichen Zweig zu besuchen. Die letzte anteilmäßig größere Gruppe innerhalb der angeführten Schulzweige ist mit 199 Personen (18,3%) dem musischen Zweig zuzuordnen. Die übrigen UntersuchungsteilnehmerInnen besuchten Schulzweige mit sportlichem (8,9 %), bildnerischen (6,2 %), darstellende Geometrie (4,2 %), wirtschaftlichem (2,5 %), humanistischen (1,8 %), Informatik (1,8 %) oder technischem (0,1 %) - Schwerpunkt. Insgesamt 68 (6,3 %) der Untersuchten führten einen anderen als die genannten Schulzweige an und 7 UntersuchungsteilnehmerInnen (0,6 %) enthielten sich der Angabe.

Tabelle 16: Häufigkeiten und Anteilswerte der besuchten Schulzweige

Schulzweige				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	Naturwissenschaftlich	247	22,6	22,7
	Sprachlicher Zweig	288	26,3	26,5
	Darstellende Geometrie	46	4,2	4,2
	Bildnerischer Zweig	67	6,1	6,2
	Musikalischer Zweig	199	18,2	18,3
	Technischer Zweig	1	0,1	0,1
	Wirtschaftlicher Zweig	27	2,5	2,5
	Humanistischer Zweig	20	1,8	1,8
	Informatikzweig	20	1,8	1,8
	Sportlicher Zweig	97	8,9	8,9
	Sonstiger Schulzweig	68	6,2	6,3
	Fehlende Angaben	7	0,6	0,6
Gesamt		1087	99,5	100,0
Fehlend	System	6	0,5	
Gesamt		1093	100,0	

Die Prüfung der Gleichverteilung der SchülerInnen auf die Schulzweige ergibt mit $p \leq .0001$ ein signifikantes Ergebnis, was bedeutet, dass keine Gleichverteilung der SchülerInnen auf die Schulzweige angenommen werden kann (siehe Tabelle 17).

Tabelle 17: Signifikanzprüfung Schulzweige

Chi ² -Test: Schulzweige	
	Schulzweige
Chi-Quadrat	1184,621 ^a
df	11
Asymptotische Signifikanz	$p \leq .0001$

7.2.4 Wunsch, Medizin zu studieren und Vorbereitungen für EMS

Im Rahmen der Gesamtstichprobe (N = 1093) können aus 1068 gültigen Fällen insgesamt 187 SchülerInnen (17,5 %) ermittelt werden, die angeben, später ein Medizinstudium anzustreben. Im Gegensatz dazu äußerte der Großteil der TestteilnehmerInnen (881, 82,5 %) keinen Wunsch, ein zukünftiges Medizinstudium absolvieren zu wollen (siehe Abbildung 10)

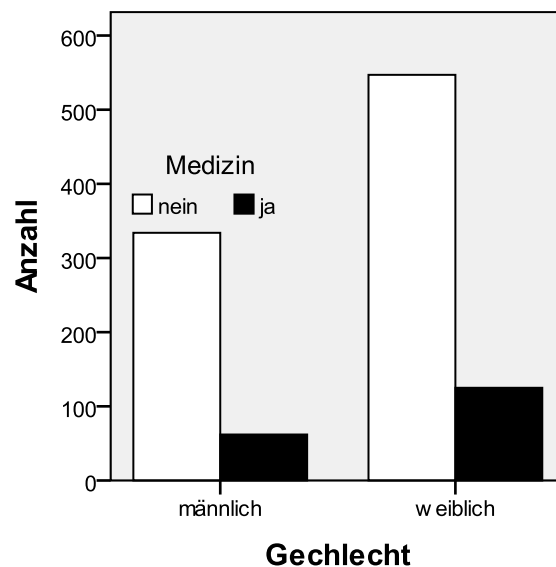


Abbildung 10: Wunsch Medizinstudium nach dem Geschlecht

Aus Tabelle 18 können die Häufigkeiten der Geschlechter nach dem Wunsch Medizin zu studieren, entnommen werden.

Tabelle 18: Häufigkeiten Geschlecht nach Wunsch Medizinstudium

Medizin			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	881	534,0	347,0
ja	187	534,0	-347,0
Gesamt	1068		

Die Gleichverteilungsprüfung ergibt mit $p \leq .0001$ ein signifikantes Ergebnis, was bedeutet, dass keine Gleichverteilung der Geschlechter in Bezug auf den Wunsch nach einem Medizinstudium, angenommen werden kann (siehe Tabelle 19).

Tabelle 19: Signifikanzprüfung Wunsch nach Medizinstudium

Chi ² -Test	
	Medizin
Chi-Quadrat	450,970 ^a
df	1
Asymptotische Signifikanz	p≤.0001

In der folgenden Kreuztabelle 20 wird im Speziellen der Wunsch nach einem Medizinstudium in Abhängigkeit vom Geschlecht veranschaulicht werden. Es zeigt sich zunächst, dass unter jenen SchülerInnen, die ein Medizinstudium beginnen wollen, anteilmäßig mehr Mädchen (66,8 %) diesen Wunsch angeben,

Tabelle 20: Kontingenztafel Geschlecht und Wunsch Medizinstudium

Kreuztabelle – Geschlechter * Wunsch nach Medizinstudium				
Geschlechter		Wunsch Medizinstudium		Gesamt
		nein	ja	
männlich	Anzahl	334	62	396
	Erwartete Anzahl	326,7	69,3	396,0
	% innerhalb von Geschlecht	84,3%	15,7%	100,0%
	Standardisierte Residuen	,4	-,9	
weiblich	Anzahl	547	125	672
	Erwartete Anzahl	554,3	117,7	672,0
	% innerhalb von Geschlecht	81,4%	18,6%	100,0%
	Standardisierte Residuen	-,3	,7	
Gesamt	Anzahl	881	187	1068
	Erwartete Anzahl	881,0	187,0	1068,0
	% innerhalb von Geschlecht	82,5%	17,5%	100,0%

Die Berechnung der Prüfgröße aus der Gesamtstichprobe ergibt mit $\chi^2(1) = 1.496$, $p = .221$ ein nicht signifikantes Ergebnis. Es kann somit kein signifikanter Verteilungsunterschied im Wunsch Medizin zu studieren in Abhängigkeit vom Geschlecht angenommen werden.

Des Weiteren ist von Interesse, ob in Abhängigkeit vom Wunsch, Medizin zu studieren bereits eine Beschäftigung mit dem EMS stattgefunden hat. Aus Tabelle 21 (Kontingenztafel) kann entnommen werden, wie viele UntersuchungsteilnehmerInnen sich im Vorfeld mit dem EMS befasst haben und gleichzeitig auch den Wunsch angeben, Medizin studieren zu wollen. In der Stichprobe von 1066 gültigen Fällen geben 117 Testpersonen an, sich bereits mit dem EMS-Test befasst zu haben, während 70 UntersuchungsteilnehmerInnen noch keine EMS- Vorbereitung anführen. 80 SchülerInnen, die kein Medizinstudium anstreben, geben an sich bereits mit dem EMS beschäftigt zu haben. Insgesamt 799 der getesteten SchülerInnen planen kein Medizinstudium zu absolvieren und haben sich bisher auch nicht mit dem EMS auseinandergesetzt.

Tabelle 21: Kontingenztafel nach Wunsch Medizinstudium und EMS-Vorbereitung

Kreuztabelle – Wunsch nach Medizinstudium * EMS-Vorbereitung					
			EMS		Gesamt
			nein	ja	
Medizin	nein	Anzahl	799	80	879
		Erwartete Anzahl	716,6	162,4	879,0
		% innerhalb von Medizin	90,9%	9,1%	100,0%
		Standardisierte Residuen	3,1	-6,5	
	ja	Anzahl	70	117	187
		Erwartete Anzahl	152,4	34,6	187,0
		% innerhalb von Medizin	37,4%	62,6%	100,0%
		Standardisierte Residuen	-6,7	14,0	
Gesamt		Anzahl	869	197	1066
		Erwartete Anzahl	869,0	197,0	1066,0
		% innerhalb von Medizin	81,5%	18,5%	100,0%

Die Berechnung der Prüfgröße $\chi^2 (1) = 292,584$, $p \leq .0001$ ergibt ein signifikantes Ergebnis. Es kann ein signifikanter Verteilungsunterschied in der Beschäftigung mit dem EMS in Abhängigkeit vom Wunsch Medizin zu studieren, angenommen werden.

Kapitel 8 Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse der Testung deskriptiv dargestellt.

8.1 Überprüfung der experimentellen Variation

Eingangs werden Reliabilitätsanalysen der EMS-ähnlichen Aufgaben und des GDT angeführt. Im Anschluss daran wird die Auslösewirksamkeit der möglichen Aktivierung von Stereotype Threat mittels des modifizierten Wissenstests in Anlehnung an das *Beliefs About the Cognitive Abilities of Females and Males Questionnaire* von Halpern und Tan (2001) dargestellt. Abschließend wird auf die Ergebnisdarstellung der geprüften Haupt- und Nebenhypothesen eingegangen.

Die Berechnungen wurden mit dem Statistik-Softwarepaket IBM Statistics SPSS 19.0 durchgeführt.

8.1.1 Reliabilitätsanalysen

Im folgenden Abschnitt werden die Ergebnisse der Reliabilitätsanalysen für die EMS-ähnlichen Aufgaben und den GDT dargestellt. Die Reliabilität (Messgenauigkeit/Zuverlässigkeit) ist das zweite Hauptgütekriterium und gibt an mit welcher Genauigkeit das Messinstrument ein bestimmtes Konstrukt misst (Bortz & Döring, 2006, S.196). Zur Bestimmung der Reliabilität wird die innere Konsistenz berechnet. Mittels Cronbach`s Alpha erfolgt eine Abschätzung der unteren Schranke der Reliabilität (Bortz & Döring, 2006, S.198).

8.1.1.1 EMS-ähnliche Aufgaben

In der folgenden Tabelle 22 werden die berechneten Reliabilitäten der sechs EMS-ähnlichen Aufgaben angegeben.

Tabelle 22: Reliabilitäten nach Cronbachs Alpha der sechs EMS-Untertests

EMS-Untertests N = 1093	Reliabilitäten (Cronbachs Alpha)
Quantitative und formale Probleme	.335
Schlauchfiguren	.65
Medizinisch-nat.wissenschaftl. GV	.353
Diagramme und Tabellen	.314
Figuren lernen	.64
Fakten lernen	.49

Die Ergebnisse der Reliabilitätsanalysen nach Cronbachs Alpha (α), für die sechs EMS-ähnlichen Aufgabenblöcke, zeigen bis auf die zwei Untertests *Schlauchfiguren* und *Figuren lernen* geringe und wenig zufriedenstellende Reliabilitäten. Es ist anzumerken, dass die Items der restlichen vier Untertests heterogen sind und somit nicht nur eine Dimension messen und erfassen, sondern mit anderen Fähigkeiten konfundieren und infolgedessen als mehrdimensional anzusehen sind. Die Trennschärfe der Items ist bis auf die zwei genannten Untertests zu niedrig. Bezüglich der Itemschwierigkeiten ist zu sagen, dass im Untertest *Quantitative und formale Probleme* jedes Item als schwierig einzustufen ist. Im Untertest *Medizinisch-naturwissenschaftliches Grundverständnis* ist die Itemschwierigkeit ab der Hälfte der Aufgaben sehr schwierig und im Untertest *Diagramme und Tabellen* zeigt sich ebenfalls, dass die Items heterogen sind und die Skalierung nicht erfüllt ist, was bedeutet, dass eine Person, die einen hohen Wert erreicht nicht automatisch gut ist. Es scheint, dass in diesem Fall mehr die Tagesverfassung der Testperson erfasst wird als die eigentliche Fähigkeit.

Für eine genauere und differenzierte Betrachtung der Reliabilitäten werden in weiterer Folge die Reliabilitäten für die Faktorstufenkombinationen der sechs EMS-ähnlichen Untertests dargestellt.

Aus Tabelle 23 können die Reliabilitäten der Faktorstufenkombinationen des EmS-ähnlichen Untertests *Quantitative und formale Probleme* entnommen werden.

Tabelle 23: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Untertest Quantitative und formale Probleme

Risiko/Stereo	A	B	C	D	Gesamt
Raten	.298 (93; 16)	.244 (90; 16)	.400 (104; 15)	.098 (86; 15)	.287(373; 16)
Ohne Instruktion	.359 (108; 16)	.296 (87; 15)	.418 (92; 16)	.297 (81; 14)	.352(368; 16)
Sicher	.379 (92; 16)	.432 (80; 10)	.213(84; 14)	.369 (96; 15)	.348(352; 16)
Gesamt	.353 (293; 16)	.343 (257; 16)	.356 (280; 16)	.279 (263; 16)	.335 (1093; 16)

In gewissen Faktorstufenkombinationen wurden einige Items von allen Personen in dieser Zelle nicht gelöst. Die Neutral-Bedingung und die Risikoinstruktion Sicher weisen mit $\alpha = .432$ die höchste Reliabilität auf.

In Tabelle 24 werden die Reliabilitäten der Faktorstufenkombinationen des EMS-ähnlichen Untertests *Schlauchfiguren* dargestellt.

Tabelle 24 Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Schlauchfiguren

Risiko/Stereo	A	B	C	D	Gesamt
Raten	.728 (93; 21)	.708 (90; 21)	.663 (104; 21)	.596 (86; 21)	.681 (373; 21)
Ohne Instruktion	.616 (108;21)	.660 (87; 21)	.712 (92; 21)	.485 (81; 21)	.641 (368; 21)
Sicher	.655 (92; 21)	.559 (80; 21)	.498 (84; 21)	.707 (96; 21)	.626 (352; 21)
Gesamt	.672 (293;21)	.657 (257;21)	.645 (280; 21)	.622 (263;21)	.651 (1093;21)

Im Untertest *Schlauchfiguren* wird in der Aktivierungs-Bedingung und in der Risikobedingung Raten mit $\alpha = .728$ die höchste Reliabilität erreicht.

Tabelle 25 zeigt die Reliabilitäten der Faktorstufenkombinationen des EMS-ähnlichen Untertests *Medizinisch-naturwissenschaftliches Grundverständnis*.

Tabelle 25 Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Untertest Medizinisch-naturwissenschaftliches Grundverständnis

Risiko/Stereo	A	B	C	D	Gesamt
Raten	.247 (93; 16)	.442 (90; 16)	.398 (104; 16)	.295 (86; 16)	.362 (373; 16)
Ohne Instruktion	.241 (108;16)	.476 (87; 16)	.380 (92; 16)	.266 (81; 15)	.352 (368; 16)
Sicher	.174 (92; 15)	.441 (80; 14)	.319 (84; 16)	.364 (96; 16)	.329 (352; 16)
Gesamt	.219 (293;16)	.467 (257; 16)	.376 (280; 16)	.317 (263;16)	.353 (1093;16)

Bei diesem Untertest zeigt sich, dass in der Neutral-Bedingung und in der Risikobedingung ohne Instruktion mit $\alpha=.476$ die höchste Reliabilität erreicht wird.

In Tabelle 26 werden die Reliabilitäten der Faktorstufenkombinationen des EMS-ähnlichen Untertests *Diagramme und Tabellen* veranschaulicht.

Tabelle 26: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Diagramme und Tabellen

Risiko/Stereo	A	B	C	D	Gesamt
Raten	.370 (93, 16)	.341 (90; 15)	.288 (104; 16)	.010 (86; 16)	.289 (373; 16)
Ohne Instruktion	.270 (108,16)	.256 (87; 16)	.262 (92; 16)	.140(81; 13)	.255 (368; 16)
Sicher	.366 (92; 15)	.411 (80; 15)	.434 (84; 16)	.320 (96; 16)	.381 (352; 16)
Gesamt	.337(293;16)	.364 (257; 16)	.332 (280; 16)	.203 (263;16)	.314(1093;16)

Bei dem Untertest *Diagramme und Tabellen* zeigen sich allgemein niedrige Reliabilitäten. Es kann angenommen werden, dass keine der Bedingungen eine besonderen Einfluss haben.

In Tabelle 27 sind die Reliabilitäten der Faktorstufenkombinationen des EMS-ähnlichen Untertests *Figuren lernen* dargestellt.

Tabelle 27: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemsanzahl) EMS- ähnlicher Untertest Figuren Lernen

Risiko/Stereotype	A	B	C	D	Gesamt
Raten	.577 (93, 8)	.682 (90, 8)	.413 (104,8)	.668 (86, 8)	.595 (373,8)
Ohne Instruktion	.622 (108,8)	.624 (87, 8)	.743 (92,8)	.580 (81, 8)	.655 (368,8)
Sicher	.631 (92, 8)	.658 (80, 8)	.534 (84, 8)	.752 (96, 8)	.659 (352,8)
Gesamt	.606 (293,8)	.654 (257,8)	.600 (280,8)	.682 (263,8)	.637 (1093)

Es zeigt sich, dass die Negativ-Aktivierungs-Bedingung und die Risikoinstruktion Raten mit $\alpha = .752$ die höchste Reliabilität aufweisen.

In der letzten Tabelle (28) werden die Reliabilitäten der Faktorstufenkombinationen des EMS-ähnlichen Untertests *Fakten lernen* dargestellt.

Tabelle 28: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemsanzahl) EMS-ähnlicher Untertest Fakten lernen

Risiko/Stereo	A	B	C	D	Gesamt
Raten	.399 (93;8)	.464 (90;8)	.309 (104;8)	.436 (86;8)	.399 (373;8)
Ohne Instruktion	.607 (108;8)	.557 (87;8)	.499 (92;8)	.377 (81;8)	.532 (368;8)
Sicher	.521 (92;8)	.580 (80;8)	.386 (84;8)	.584 (96;8)	.524 (352;8)
Gesamt	.521 (293;8)	.531 (257;8)	.404 (280;8)	.490 (263;8)	.489 (1093;8)

Bei dem Untertest *Fakten lernen* kann in der Aktivierungs-Bedingung und in der Risikobedingung ohne Instruktion, mit $\alpha = .607$ die höchste Reliabilität festgestellt werden.

Infolge der Berechnung der Reliabilitäten der Faktorstufenkombinationen der EMS-ähnlichen Untertests zeigt sich, dass allgemein die Reliabilitäten geringfügig erhöht werden konnten. Darüber hinaus ist anzumerken, dass sich die Reliabilitäten der verschiedenen Faktorstufenkombinationen unterscheiden und unterschiedlich hohe Reliabilitäten aufweisen. Abschließend ist zu sagen, dass die

Untertests Figuren und Fakten lernen sowie Schlauchfiguren die höchsten Reliabilitäten aufweisen.

8.1.1.2 GDT

Die Reliabilitätsanalyse nach Cronbachs Alpha ergibt für den GDT mit $\alpha = .891$ eine zufriedenstellende Reliabilität.

8.2 Auslösung der Stereotype Threat - Bedingungen

Im Rahmen der Testung wurde den UntersuchungsteilnehmerInnen, vor der Bearbeitung der EMS-ähnlichen Aufgaben, der *Wissenstest* zur Aktivierung der vier Stereotype Threat-Bedingungen vorgegeben. Die Aufgabenstellungen des Wissenstest wurden in Anlehnung an das *Beliefs About the Cognitive Abilities of Females and Males Questionnaire* von Halpern und Tan (2001) konstruiert und ausgewertet. Demzufolge werden die Mittelwerte als Kriterium der einzelnen Aufgabenstellungen herangezogen, um die Auslösung der Stereotype Threat-Bedingungen zu überprüfen. Für diese Überprüfung gilt, dass beobachtete Prozentwerte, die außerhalb des Bereichs von 45,0 bis 55,0 Prozent liegen, als erfolgreiche Aktivierung der Bedingungen *Positiv – und Negativaktivierung* gewertet werden können, während Anteilswerte, die zwischen 45,0 und 55,0 Prozent liegen den Bedingungen *Neutral* und *Aufklärung* zugeordnet werden. Zur Überprüfung der Wirksamkeit der vier Stereotype Threat-Bedingungen in den sechs EMS-Bereichen werden χ^2 -Tests durchgeführt. Mittels χ^2 - Test wird geprüft, ob sich die beobachteten Häufigkeiten der Merkmalskombinationen signifikant von den zu erwartenden Häufigkeiten unterscheiden bzw. ob sich die Daten in der Matrix signifikant von einer Gleichverteilung unterscheiden. Ein signifikantes Ergebnis gilt als Hinweis, dass es Verteilungsunterschiede in der abhängigen kategorialen Variable in Abhängigkeit von der unabhängigen Variable anzunehmen sind (Bortz & Döring, 2006, S.613).

Zusätzlich wurde im Rahmen des *Wissentests* eine Aufgabe zur Überprüfung eines Stereotyps bezüglich der Theory of Mind vorgegeben, da in der Literatur mehrmals bestätigt werden konnte, dass Frauen über eine ausgeprägtere ToM verfügen als Männer und generell gesellschaftlich als sozial kompetenter wahrgenommen und eingeschätzt werden.

Die Berechnungen ergeben folgendes Bild:

Aus Tabelle 29 können die Häufigkeiten die Häufigkeiten der Fremdeinschätzungsaufgabe *Figuren lernen* entnommen werden.

Tabelle 29: Häufigkeiten Figuren lernen N= 1084

Figuren lernen			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	288	361,3	-73,3
ja	646	361,3	284,7
entgegengesetzt	150	361,3	-211,3
Gesamt	1084		

In der folgenden Kontingenztafel (Tabelle 30) werden die vier Stereotype Threat-Bedingungen in Abhängigkeit von der Fremdeinschätzungsaufgabe Figuren lernen dargestellt.

Tabelle 30: Kontingenztafel Stereotype Threat-Bedingungen und WT Figuren lernen (N= 1084)

Kreuztabelle Stereotype * Figuren lernen						
		Figuren			Gesamt	
		nein	ja	entgegengesetzt		
Stereotype	A Aktivierung	Anzahl	115	124	50	289
		Erwartete Anzahl	76,8	172,2	40,0	289,0
		% innerhalb von Stereotype	39,8%	42,9%	17,3%	100,0%

		Standardisierte Residuen	4,4	-3,7	1,6	
B neutral	Anzahl		37	205	13	255
	Erwartete Anzahl		67,7	152,0	35,3	255,0
	% innerhalb von Stereotype		14,5%	80,4%	5,1%	100,0%
	Standardisierte Residuen		-3,7	4,3	-3,8	
C Aufklärung	Anzahl		33	216	30	279
	Erwartete Anzahl		74,1	166,3	38,6	279,0
	% innerhalb von Stereotype		11,8%	77,4%	10,8%	100,0%
	Standardisierte Residuen		-4,8	3,9	-1,4	
D Negativaktivierung	Anzahl		103	101	57	261
	Erwartete Anzahl		69,3	155,5	36,1	261,0
	% innerhalb von Stereotype		39,5%	38,7%	21,8%	100,0%
	Standardisierte Residuen		4,0	-4,4	3,5	
Gesamt	Anzahl		288	646	150	1084
	Erwartete Anzahl		288,0	646,0	150,0	1084,0
	% innerhalb von Stereotype		26,6%	59,6%	13,8%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 168.72$, $p \leq .0001$ ergibt ein signifikantes Ergebnis. Es können Verteilungsunterschiede in der Fremdeinschätzung in Abhängigkeit von der Stereotype Threat-Bedingung angenommen werden.

In Tabelle 31 werden die Häufigkeiten der Fremdeinschätzungsaufgabe *Fakten lernen* dargestellt.

Tabelle 31: Häufigkeiten WT Fakten lernen N= 1082

Fakten			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	307	360,7	-53,7
ja	580	360,7	219,3
entgegengesetzt	195	360,7	-165,7
Gesamt	1082		

Die Verteilungen der UntersuchungsteilnehmerInnen in den Stereotype Threat-Bedingungen in Abhängigkeit von der Fremdeinschätzungsaufgabe *Fakten lernen* können aus der Kontingenztafel (Tabelle 32) entnommen werden.

Tabelle 32: Kontingenztafel Stereotype Threat-Bedingungen und WT Fakten lernen (N = 1082)

Kreuztabelle Stereotype * Fakten						
			Fakten			Gesamt
			nein	ja	entgegengesetzt	
Stereotype	A Aktivierung	Anzahl	48	205	37	290
		Erwartete Anzahl	82,3	155,5	52,3	290,0
		% innerhalb von Stereotype	16,6%	70,7%	12,8%	100,0%
		Standardisierte Residuen	-3,8	4,0	-2,1	
	B neutral	Anzahl	84	132	38	254
		Erwartete Anzahl	72,1	136,2	45,8	254,0
		% innerhalb von Stereotype	33,1%	52,0%	15,0%	100,0%
		Standardisierte Residuen	1,4	-,4	-1,1	
	C Aufklärung	Anzahl	127	104	47	278
		Erwartete Anzahl	78,9	149,0	50,1	278,0
		% innerhalb von Stereotype	45,7%	37,4%	16,9%	100,0%
		Standardisierte Residuen	5,4	-3,7	-,4	

	D Negativaktivierung	Anzahl	48	139	73	260
		Erwartete Anzahl	73,8	139,4	46,9	260,0
		% innerhalb von Stereotype	18,5%	53,5%	28,1%	100,0%
		Standardisierte Residuen	-3,0	,0	3,8	
Gesamt		Anzahl	307	580	195	1082
		Erwartete Anzahl	307,0	580,0	195,0	1082,0
		% innerhalb von Stereotype	28,4%	53,6%	18,0%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 104.70$, $p \leq .0001$ ergibt ein signifikantes Ergebnis, was bedeutet, dass Verteilungsunterschiede in der Fremdeinschätzung *Fakten lernen* in Abhängigkeit von der Stereotype Threat-Bedingung angenommen werden können.

Tabelle 33 veranschaulicht die Häufigkeiten der Fremdeinschätzungsaufgabe *Medizinisch-naturwissenschaftliches Grundverständnis*.

Tabelle 33: Häufigkeiten WT Med.-naturwissenschaftl. GV N= 1080

Naturwissenschaftliches Grundverständnis			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	334	360,0	-26,0
ja	571	360,0	211,0
entgegengesetzt	175	360,0	-185,0
Gesamt	1080		

In Bezug auf die Fremdeinschätzungsaufgabe *Medizinisch-naturwissenschaftliches Grundverständnis* können die Verteilungen der SchülerInnen (N = 1080) aus der folgenden Kontingenztabelle (Tabelle 34) entnommen werden.

Tabelle 34: Kontingenztabelle Stereotype Threat-Bedingungen und WT Med.-nat.-wissenschaftl. GV

Stereotype * Natur Kreuztabelle						
			Natur			Gesamt
			nein	ja	entgegengesetzt	
Stereotype	A Aktivierung	Anzahl	100	149	41	290
		Erwartete Anzahl	89,7	153,3	47,0	290,0
		% innerhalb von Stereotype	34,5%	51,4%	14,1%	100,0%
		Standardisierte Residuen	1,1	-,3	-,9	
	B neutral	Anzahl	76	155	23	254
		Erwartete Anzahl	78,6	134,3	41,2	254,0
		% innerhalb von Stereotype	29,9%	61,0%	9,1%	100,0%
		Standardisierte Residuen	-,3	1,8	-2,8	
	C Aufklärung	Anzahl	87	150	40	277
		Erwartete Anzahl	85,7	146,5	44,9	277,0
		% innerhalb von Stereotype	31,4%	54,2%	14,4%	100,0%
		Standardisierte Residuen	,1	,3	-,7	
	D Negativaktivierung	Anzahl	71	117	71	259
		Erwartete Anzahl	80,1	136,9	42,0	259,0
		% innerhalb von Stereotype	27,4%	45,2%	27,4%	100,0%
		Standardisierte Residuen	-1,0	-1,7	4,5	
Gesamt		Anzahl	334	571	175	1080
		Erwartete Anzahl	334,0	571,0	175,0	1080,0
		% innerhalb von Stereotype	30,9%	52,9%	16,2%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 38,02, p \leq .0001$ ergibt ein signifikantes Ergebnis. Es können somit Verteilungsunterschiede in der Fremdeinschätzung, bezüglich der Aufgabe *Medizinisch-naturwissenschaftliches*

Grundverständnis in Abhängigkeit von der Stereotype Threat-Bedingung angenommen werden.

In Tabelle 35 werden die Häufigkeiten der Fremdeinschätzungsaufgabe *Diagramme und Tabellen* dargestellt.

Tabelle 35; Häufigkeiten WT Diagramme und Tabellen N= 1080

Diagramme und Tabellen			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	320	360,0	-40,0
ja	637	360,0	277,0
entgegengesetzt	123	360,0	-237,0
Gesamt	1080		

Aus der Kontingenztafel (Tabelle 36) sind die Verteilungen bezüglich der Fremdeinschätzungsaufgabe *Diagramme und Tabellen*, für 1080 gültige Fälle, ersichtlich.

Tabelle 36: Kontingenztafel Stereotype Threat-Bedingungen und WT Diagramme und Tabellen

Kreuztabelle Stereotype * Diagramme und Tabellen						
			Tabellen			Gesamt
			nein	ja	entgegengesetzt	
Stereotype	A Aktivierung	Anzahl	52	223	14	289
		Erwartete Anzahl	85,6	170,5	32,9	289,0
		% innerhalb von Stereotype	18,0%	77,2%	4,8%	100,0%
		Standardisierte Residuen	-3,6	4,0	-3,3	
	B neutral	Anzahl	95	139	20	254

		Erwartete Anzahl	75,3	149,8	28,9	254,0
		% innerhalb von Stereotype	37,4%	54,7%	7,9%	100,0%
		Standardisierte Residuen	2,3	-,9	-1,7	
	C Aufklärung	Anzahl	125	123	29	277
		Erwartete Anzahl	82,1	163,4	31,5	277,0
		% innerhalb von Stereotype	45,1%	44,4%	10,5%	100,0%
		Standardisierte Residuen	4,7	-3,2	-,5	
	D Negativaktivierung	Anzahl	48	152	60	260
		Erwartete Anzahl	77,0	153,4	29,6	260,0
		% innerhalb von Stereotype	18,5%	58,5%	23,1%	100,0%
		Standardisierte Residuen	-3,3	-,1	5,6	
Gesamt		Anzahl	320	637	123	1080
		Erwartete Anzahl	320,0	637,0	123,0	1080,0
		% innerhalb von Stereotype	29,6%	59,0%	11,4%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 123.77$, $p \leq .0001$ ergibt einen signifikanten Verteilungsunterschied in der Fremdeinschätzung in Abhängigkeit von der Stereotype Threat-Bedingung.

Aus Tabelle 37 können die Häufigkeiten der Fremdeinschätzungsaufgabe *Quantitative und formale Probleme* entnommen werden.

Tabelle 37: Häufigkeiten WT Quantitative und formale Probleme N= 1080

Quantitative und formale Probleme			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	339	360,0	-21,0
ja	590	360,0	230,0
entgegengesetzt	151	360,0	-209,0
Gesamt	1080		

Für die Fremdeinschätzungsaufgabe *Quantitative und formale Probleme* werden die Verteilungen in der folgenden Kontingenztafel (Tabelle 38) dargestellt.

Tabelle 38: Kontingenztafel Stereotype Threat-Bedingungen und WT Quantitative und formale Probleme (N = 1080)

Kreuztabelle Stereotype * Quantitative und formale Probleme						
			Formal			Gesamt
			nein	ja	entgegengesetzt	
Stereotype	A Aktivierung	Anzahl	60	207	22	289
		Erwartete Anzahl	90,7	157,9	40,4	289,0
		% innerhalb von Stereotype	20,8 %	71,6%	7,6%	100,0%
		Standardisierte Residuen	-3,2	3,9	-2,9	
	B neutral	Anzahl	86	142	26	254
		Erwartete Anzahl	79,7	138,8	35,5	254,0
		% innerhalb von Stereotype	33,9 %	55,9%	10,2%	100,0%
		Standardisierte Residuen	,7	,3	-1,6	
	C Aufklärung	Anzahl	141	114	22	277
		Erwartete Anzahl	86,9	151,3	38,7	277,0
		% innerhalb von Stereotype	50,9 %	41,2%	7,9%	100,0%
		Standardisierte Residuen	5,8	-3,0	-2,7	
	D	Anzahl	52	127	81	260

	Negativaktivierung	Erwartete Anzahl	81,6	142,0	36,4	260,0
		% innerhalb von Stereotype	20,0 %	48,8%	31,2%	100,0%
		Standardisierte Residuen	-3,3	-1,3	7,4	
Gesamt		Anzahl	339	590	151	1080
		Erwartete Anzahl	339,0	590,0	151,0	1080,0
		% innerhalb von Stereotype	31,4 %	54,6%	14,0%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 154.39 =, p \leq .0001$ ergibt einen signifikanten Verteilungsunterschied in der Fremdeinschätzung *Quantitative und formale Probleme*, in Abhängigkeit von der Stereotype Threat-Bedingung.

In Tabelle 39 werden die Häufigkeiten der Fremdeinschätzungsaufgabe *Schlauchfiguren* dargestellt.

Tabelle 39: Häufigkeiten WT Schlauchfiguren N=1080

Schlauchfiguren			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	314	360,0	-46,0
ja	651	360,0	291,0
entgegengesetzt	115	360,0	-245,0
Gesamt	1080		

Abschließend werden die Verteilungen der UntersuchungsteilnehmerInnen in der Fremdeinschätzungsaufgabe Schlauchfiguren dargestellt (siehe Tabelle 40).

**Tabelle 40: Kontingenztabelle Stereotype Threat-Bedingungen und WT
Quantitative und formale Probleme (N = 1080)**

Kreuztabelle Stereotype * Schlauchfiguren						
			Schlauchfiguren			Gesamt
			nein	ja	entgegengesetzt	
Stereotype	A Aktivierung	Anzahl	70	210	9	289
		Erwartete Anzahl	84,0	174,2	30,8	289,0
		% innerhalb von Stereotype	24,2%	72,7%	3,1%	100,0%
		Standardisierte Residuen	-1,5	2,7	-3,9	
	B neutral	Anzahl	77	149	28	254
		Erwartete Anzahl	73,8	153,1	27,0	254,0
		% innerhalb von Stereotype	30,3%	58,7%	11,0%	100,0%
		Standardisierte Residuen	,4	-,3	,2	
	C Aufklärung	Anzahl	102	147	28	277
		Erwartete Anzahl	80,5	167,0	29,5	277,0
		% innerhalb von Stereotype	36,8%	53,1%	10,1%	100,0%
		Standardisierte Residuen	2,4	-1,5	-,3	
	D Negativaktivierung	Anzahl	65	145	50	260
		Erwartete Anzahl	75,6	156,7	27,7	260,0
		% innerhalb von Stereotype	25,0%	55,8%	19,2%	100,0%
		Standardisierte Residuen	-1,2	-,9	4,2	
Gesamt		Anzahl	314	651	115	1080
		Erwartete Anzahl	314,0	651,0	115,0	1080,0
		% innerhalb von Stereotype	29,1%	60,3%	10,6%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 53.91 =$, $p \leq .0001$ ergibt ein signifikantes Ergebnis. Es können Verteilungsunterschiede in der Fremdeinschätzung, in der Aufgabe *Schlauchfiguren*, in Abhängigkeit von der Stereotype Threat-Bedingung angenommen werden. Das bedeutet, dass ein

Zusammenhang zwischen der Fremdeinschätzung und der Stereotype Threat-Bedingung beobachtet werden kann.

In Tabelle 41 werden die Signifikanzprüfungen der Fremdeinschätzungsaufgaben WT dargestellt.

Tabelle 41: Signifikanzbeurteilungen der Fremdeinschätzungsaufgaben

Statistik für Test						
	Fig	Fakten	Natur	Tabellen	Formal	Schlauch
$\chi^2 (2)$	362,753 ^a	217,466 ^b	220,617 ^c	373,606 ^c	269,506 ^c	407,839 ^c
Asymptotische Signifikanz	$p \leq .0001$	$p \leq .0001$	$p \leq .0001$	$p \leq .0001$	$p \leq .0001$	$p \leq .0001$

Die Wirksamkeit der Stereotype Threat-Bedingungen kann an den bereits unterschiedlichen Prozentmittelwerten beurteilt werden. Des Weiteren erfolgt die Darstellung der Ergebnisse unter Berücksichtigung des Geschlechts und geschlechtsspezifischer Unterschiede.

Aktivierung:

In Bezug auf die erste Bedingung *Aktivierung* kann eine Auslösung des angestrebten Stereotyps angenommen werden. Es zeigt sich, dass die Ergebnisse mit den Befunden der vorliegenden Literatur konsistent sind und Mädchen, wie beschrieben, in der Aktivierungsbedingung bessere Fähigkeiten im Gedächtnis (*Figuren lernen*: $M = 57.25$, $SD = 24.12$; *Fakten lernen*: $M = 66.22$, $SD = 23.20$) und in der *ToM* ($M = 78.36$, $SD = 17.08$) zugesprochen werden. Die Stärken der Burschen hingegen werden, wie beschrieben, in den Aufgaben *Quantitative und formale Probleme* ($M = 68.56$, $SD = 21.21$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($M = 59.21$, $SD = 22.24$) und *Diagramme und Tabellen* ($M = 68.71$, $SD = 19.78$) gesehen (siehe Tabelle 42).

Tabelle 42 Statistische Kennwerte der Aktivierungs-Bedingung (A)

Stereotype		Statistiken														
		W1 M	W1 F	W2 M	W2 F	W3 M	W3 F	W4 M	W4 F	W5 M	W5 F	W6 M	W6 F	W7 M	W7 F	
A Aktivierung	N	Gültig	289	289	290	290	290	290	289	289	289	289	289	289	288	288
		Fehlend	4	4	3	3	3	3	4	4	4	4	4	4	5	5
	M		42,7	57,25	33,69	66,22	59,21	40,76	69,71	30,35	68,56	31,54	71,65	28,45	21,71	78,36
		SD	24,143	24,122	23,036	23,196	22,244	22,257	19,777	19,829	21,207	21,322	20,886	21,023	17,165	17,078

Neutral:

In dieser Bedingung kann, wie erwartet, in den Einschätzungen der UntersuchungsteilnehmerInnen keines der gewünschten Stereotype beobachtet werden. Es zeigt sich einzig in der Aufgabe zur Einschätzung der *ToM*, dass sowohl Burschen als auch Mädchen, den Frauen eine bessere und ausgeprägtere *ToM* ($M = 65.99$, $SD = 14.40$) zuschreiben (siehe Tabelle 43).

Tabelle 43:: Statistische Kennwerte der Neutral-Bedingung (B)

B neutral	N	Gültig	255	255	254	254	254	254	254	254	254	254	254	254	254	254
		Fehlend	2	2	3	3	3	3	3	3	3	3	3	3	3	3
	M		49,47	50,76	48,13	51,87	52,78	47,26	54,57	45,44	53,70	46,10	53,08	47,00	34,25	65,99
		SD	8,000	7,981	13,263	13,263	9,611	9,585	9,236	9,238	12,333	11,863	12,247	12,137	14,394	14,404

Aufklärung:

In der dritten Bedingung *Aufklärung* können keine bis geringe Effekte bezüglich der Stereotype beobachtet werden. Es zeigt sich wiederholt, dass die getesteten SchülerInnen, Frauen eine bessere Fähigkeit in der *ToM* ($M = 68.32$, $SD = 13.80$) zuschreiben. Des Weiteren werden Männern bessere Leistungen in den Aufgaben

Quantitative und formale Probleme ($M = 56.45$, $SD = 11.80$) und *Diagramme und Tabellen* ($M = 55.09$, $SD = 11.51$, siehe Tabelle 44).

Tabelle 44: Statistische Kennwerte der Aufklärungs-Bedingung (C)

C Aufklär ung	N	Gültig	279	279	278	278	277	277	277	277	277	277	277	276	276
		Fehlend	1	1	2	2	3	3	3	3	3	3	3	3	4
	M	50,0	49,90	46,98	53,00	52,12	47,88	55,09	44,84	56,45	43,58	54,62	45,36	31,68	68,32
	SD	8,60	8,730	14,125	14,221	9,706	9,668	11,511	11,142	11,803	11,810	12,270	12,274	13,803	13,803

Negativ-Aktivierung:

In der vierten und letzten Bedingung kann eine erfolgreiche Auslösung der gewünschten Stereotype beobachtet werden. Es zeigt sich die angestrebte Umkehrung der Stereotype, in dem Frauen eine Überlegenheit in den vier Bereichen *Quantitative und formale Probleme* ($M = 57.06$, $SD = 26.38$), *Schlauchfiguren* ($M = 63.89$, $SD = 26.50$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($M = 58.22$, $SD = 25.09$) und *Diagramme und Tabellen* ($M = 64.12$, $SD = 24.79$) zugeschrieben, während Männer in den Gedächtnis-Aufgaben *Figuren lernen* ($M = 55.95$, $SD = 25.29$) und *Fakten lernen* ($M = 61.13$, $SD = 26.79$) besser eingeschätzt werden. Einzig in der *ToM*-Aufgabe bleibt die Einschätzung der UntersuchungsteilnehmerInnen konsistent, in dem Frauen erneut bessere *ToM*-Fähigkeiten ($M = 77.82$, $SD = 18,74$) zugesprochen werden (siehe Tabelle 45).

Tabelle 45: Statistische Kennwerte der Negativ-Aktivierungs-Bedingung (D)

D Negativ aktivier ung	N	Gültig	261	261	260	260	259	260	260	260	260	260	260	260	257	257
		Fehlend	2	2	3	3	4	3	3	3	3	3	3	3	3	6
	M	55,9	44,09	61,13	38,60	41,60	58,22	36,09	64,12	42,84	57,06	36,35	63,89	22,19	77,82	
	SD	25.9	253	26.9	26,3	24,5	25,9	25,2	24,9	26,6	26,8	26,7	26,5	18,7	18,7	

Es zeigt sich, dass unabhängig von den vier Stereotype Threat Frauen höhere Werte in der Empathiefähigkeit und sozialen Kompetenz zugeschrieben werden. Die aus den bisherigen Forschungen bekannten Unterschiede zugunsten der Frauen werden auch in dieser Untersuchung bestätigt, da Frauen von allen Untersuchungsteilnehmern als empathischer und sozial kompetenter eingestuft werden.

Selbstratings:

Die Selbstratings des Wissenstests werden mittels zweifaktorieller multivariater Varianzanalyse in Abhängigkeit von den Geschlechtern und den unterschiedlichen Stereotype überprüft. Die statistischen Kennwerte sind Tabelle 46 zu entnehmen.

Tabelle 46: Statistische Kennwerte der Selbsteinschätzungsaufgaben

Deskriptive Statistiken					
	Geschlecht	Stereotype	M	SD	N
Wissen8	männlich	A Aktivierung	6,59	,775	119
		B neutral	6,73	,664	83
		C Aufklärung	6,55	1,106	108
		D Negativaktivierung	6,67	,836	85
		Gesamt	6,63	,870	395
	weiblich	A Aktivierung	6,55	1,049	168
		B neutral	6,51	1,000	164
		C Aufklärung	6,53	1,033	165
		D Negativaktivierung	6,41	1,083	173
		Gesamt	6,50	1,041	670
	Gesamt	A Aktivierung	6,56	,943	287
		B neutral	6,58	,906	247
		C Aufklärung	6,54	1,060	273
		D Negativaktivierung	6,50	1,014	258
		Gesamt	6,55	,983	1065
Wissen9	männlich	A Aktivierung	4,84	1,275	119
		B neutral	5,16	1,418	83
		C Aufklärung	4,81	1,284	108
		D Negativaktivierung	4,98	1,215	85
		Gesamt	4,93	1,298	395
	weiblich	A Aktivierung	5,05	1,210	168

		B neutral	5,17	1,170	164
		C Aufklärung	5,10	1,213	165
		D Negativaktivierung	4,88	1,293	173
		Gesamt	5,05	1,225	670
	Gesamt	A Aktivierung	4,97	1,240	287
		B neutral	5,17	1,256	247
		C Aufklärung	4,99	1,247	273
		D Negativaktivierung	4,91	1,266	258
		Gesamt	5,01	1,254	1065
Wissen10	männlich	A Aktivierung	5,30	1,344	119
		B neutral	5,10	1,385	83
		C Aufklärung	5,05	1,469	108
		D Negativaktivierung	4,84	1,511	85
		Gesamt	5,09	1,428	395
	weiblich	A Aktivierung	4,82	1,445	168
		B neutral	4,93	1,451	164
		C Aufklärung	4,82	1,402	165
		D Negativaktivierung	4,89	1,260	173
		Gesamt	4,86	1,388	670
	Gesamt	A Aktivierung	5,02	1,421	287
		B neutral	4,98	1,428	247
		C Aufklärung	4,91	1,431	273
		D Negativaktivierung	4,87	1,345	258
		Gesamt	4,95	1,407	1065
Wissen11	männlich	A Aktivierung	5,26	1,498	119
		B neutral	5,42	1,458	83
		C Aufklärung	5,07	1,569	108
		D Negativaktivierung	4,56	1,644	85
		Gesamt	5,09	1,565	395
	weiblich	A Aktivierung	3,96	1,615	168
		B neutral	4,23	1,706	164
		C Aufklärung	4,24	1,418	165
		D Negativaktivierung	4,45	1,612	173
		Gesamt	4,22	1,597	670
	Gesamt	A Aktivierung	4,50	1,691	287
		B neutral	4,63	1,720	247
		C Aufklärung	4,57	1,533	273
		D Negativaktivierung	4,48	1,620	258
		Gesamt	4,54	1,640	1065

Wissen12	männlich	A Aktivierung	5,43	1,499	119
		B neutral	5,00	1,753	83
		C Aufklärung	4,93	1,700	108
		D Negativaktivierung	4,74	1,767	85
		Gesamt	5,05	1,683	395
	weiblich	A Aktivierung	3,99	1,771	168
		B neutral	4,06	1,846	164
		C Aufklärung	3,98	1,737	165
		D Negativaktivierung	4,12	1,801	173
		Gesamt	4,04	1,786	670
	Gesamt	A Aktivierung	4,59	1,807	287
		B neutral	4,38	1,865	247
		C Aufklärung	4,36	1,781	273
		D Negativaktivierung	4,32	1,810	258
		Gesamt	4,41	1,815	1065
Wissen13	männlich	A Aktivierung	6,07	1,079	119
		B neutral	5,90	1,376	83
		C Aufklärung	5,68	1,570	108
		D Negativaktivierung	5,89	1,581	85
		Gesamt	5,89	1,403	395
	weiblich	A Aktivierung	5,04	1,786	168
		B neutral	5,21	1,664	164
		C Aufklärung	5,24	1,649	165
		D Negativaktivierung	5,17	1,833	173
		Gesamt	5,16	1,734	670
	Gesamt	A Aktivierung	5,47	1,612	287
		B neutral	5,45	1,604	247
		C Aufklärung	5,41	1,629	273
		D Negativaktivierung	5,41	1,784	258
		Gesamt	5,43	1,656	1065
Wissen14	männlich	A Aktivierung	4,93	1,566	119
		B neutral	5,36	1,597	83
		C Aufklärung	5,46	1,307	108
		D Negativaktivierung	5,42	1,366	85
		Gesamt	5,27	1,476	395
	weiblich	A Aktivierung	6,42	,970	168
		B neutral	6,28	,994	164
		C Aufklärung	6,39	,922	165
		D Negativaktivierung	6,42	,909	173

		Gesamt	6,38	,949	670
	Gesamt	A Aktivierung	5,80	1,450	287
		B neutral	5,97	1,302	247
		C Aufklärung	6,03	1,180	273
		D Negativaktivierung	6,09	1,177	258
		Gesamt	5,97	1,288	1065

In die Berechnungen werden 1065 gültige Fälle eingeschlossen und es zeigt sich in Abhängigkeit von Geschlecht und unabhängig von den Stereotype Threat-Bedingungen, dass sich Burschen in fünf Bereichen *Quantitative und formale Probleme* $F(1, 1065) = 78.300, p \leq .0001$, *Schlauchfiguren* $F(1, 1065) = 48.401, p \leq .001$, *Medizinisch-naturwissenschaftliches Grundverständnis* $F(1, 1065) = 5.277, p = .022$, *Figuren lernen* $F(1, 1065) = 4.677, p = .031$ und *Diagramme und Tabellen* $F(1, 1065) = 73.549, p \leq .0001$ signifikant besser einschätzen als Mädchen. Einzig in der Gedächtnisaufgabe *Fakten lernen* $F(1, 1065) = 1.753, p = .186$ können keine signifikanten geschlechtsspezifischen Unterschiede in den Selbsteinschätzungen beobachtet werden (siehe Tabelle 47).

Tabelle 47: : Selbsteinschätzungen WT

	Geschlecht	M	SD	F(1,1057)	Sig
Quant. und formale Probleme	männlich	5.05	1.68	78.30	≤ .0001
	weiblich	4.04	1.79		
Schlauchfiguren	männlich	5.89	1.40	48.40	≤ .0001
	weiblich	5.16	1.73		
Med.-naturwissenschaftl. Grundverständnis	männlich	5.09	1.43	5.28	.022
	weiblich	4.86	1.39		
Diagramme und Tabellen	männlich	5.09	1.57	73.55	≤ .0001
	weiblich	4.22	1.60		

Figuren lernen	männlich	6.63	0.87	4.67	.031
	weiblich	6.50	1.04		
Fakten lernen	männlich	4.93	1.30	1.75	.186
	weiblich	5.05	1.23		

Aus Tabelle 48 können die statistischen Kennwerte der ToM – Selbsteinschätzungsaufgabe entnommen werden.

Tabelle 48: Statistische Kennwerte ToM-Selbsteinschätzungsaufgabe

Deskriptive Statistiken				
Abhängige Variable:Wissen14- ToM				
Geschlecht	Stereotype	<i>M</i>	<i>SD</i>	N
männlich	A Aktivierung	4,93	1.56	120
	B neutral	5,29	1.64	85
	C Aufklärung	5,46	1.31	108
	D Negativaktivierung	5,41	1.37	86
	Gesamt	5,26	1.48	399
weiblich	A Aktivierung	6,43	.97	169
	B neutral	6,28	.99	166
	C Aufklärung	6,39	.92	166
	D Negativaktivierung	6,42	.91	173
	Gesamt	6,38	.95	674
Gesamt	A Aktivierung	5,81	1.45	289
	B neutral	5,94	1.33	251
	C Aufklärung	6,03	1.18	274
	D Negativaktivierung	6,08	1.18	259
	Gesamt	5,96	1.30	1073

Bei den Mädchen zeigt sich, dass sie sich in der *ToM* $F(1, 1065) = 211.655, p \leq .0001$ besser einschätzen

Im Zuge der Überprüfung der Stereotype Threat-Bedingungen unabhängig vom Geschlecht zeigt sich, dass nur bei der *ToM-Aufgabe* Selbsteinschätzung mit der entsprechenden Prüfgröße $F(3, 1057) = 2.69, p = .045$ ein signifikanter Unterschied beobachtet werden kann (siehe Tabelle 49).

Tabelle 49: ANOVA ToM Selbsteinschätzungsaufgabe

Tests der Zwischensubjekteffekte					
Abhängige Variable:Wissen14					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	338,389 ^a	7	48,341	35,286	$p \leq .0001$
Konstanter Term	33567,755	1	33567,755	24502,102	$p \leq .0001$
Geschlecht	301,712	1	301,712	220,228	$p \leq .0001$
Stereotype	10,847	3	3,616	2,639	$p = .048$
Geschlecht* Stereotype	13,833	3	4,611	3,366	$p \leq .018$
Fehler	1459,045	1065	1,370		
Gesamt	39935,000	1073			
Korrigierte Gesamtvariation	1797,433	1072			

Die Berechnung eines Post-hoc-Tests nach Bonferroni zur Prüfung der paarweisen Unterschiede zwischen den Stereotype Threat-Bedingungen ergibt nur zwischen Bedingung Aktivierung ($M = 5.81, SD = 1.45, 289$ gültige Fälle) und der Bedingung Negativaktivierung ($M = 6.08, SD = 1.18, 259$ gültige Fälle) mit $p = .033$ ein signifikantes Ergebnis (siehe Tabelle 50). Die Aktivierungsbedingung hat einen mindernden Einfluss auf die Selbsteinschätzung in der ToM. Die Negativaktivierung erreicht die signifikant höhere ToM-Selbsteinschätzung.

Tabelle 50: Post hoc-Tests nach Bonferroni

Wissen14-ToM						
Bonferroni						
(I) Stereotype	(J) Stereotype	Mittlere Differenz (I-J)	Standardfehler	Sig.	95%-Konfidenzintervall	
					Untergrenze	Obergrenze
A Aktivierung	B neutral	-,14	,101	$p=1.000$	-,40	,13
	C Aufklärung	-,22	,099	$p=.159$	-,48	,04
	D Negativaktivierung	-,28*	,100	$p=.033$	-,54	-,01
B neutral	A Aktivierung	,14	,101	$p=1.000$	-,13	,40
	C Aufklärung	-,08	,102	$p=1.000$	-,35	,19
	D Negativaktivierung	-,14	,104	$p=1.000$	-,41	,13
C Aufklärung	A Aktivierung	,22	,099	$p=.159$	-,04	,48
	B neutral	,08	,102	$p=1.000$	-,19	,35
	D Negativaktivierung	-,06	,101	$p=1.000$	-,33	,21
D Negativaktivierung	A Aktivierung	,28*	,100	$p=.033$,01	,54
	B neutral	,14	,104	$p=1.000$	-,13	,41
	C Aufklärung	,06	,101	$p=1.000$	-,21	,33

In den Bereichen *Diagramme und Tabellen* $F(3, 1065) = 6.820, p \leq .0001$ und *ToM* $F(3, 1065) = 3.637, p = .013$ können signifikante Interaktionen in Abhängigkeit vom Geschlecht und den Stereotype Threat- Bedingungen beobachtet werden (siehe Abbildungen 11 und 12).

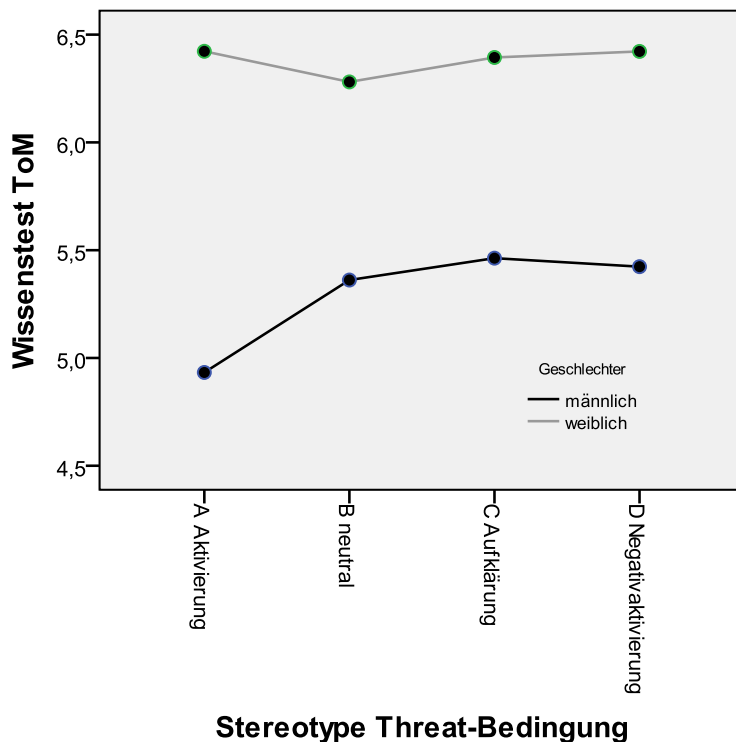


Abbildung 11: Interaktionsdiagramm ToM-Aufgabe

Aus Abbildung 11 ist ersichtlich, dass sich Burschen bei den *ToM-Aufgaben*, in der Bedingung *Aktivierung* schlechter einschätzen als Mädchen. Die Berechnung der entsprechenden Prüfgröße ergibt mit $F(1, 1057) = 211.655, p \leq .0001$. Es kann ein geschlechtsspezifischer Unterschied zugunsten der Mädchen angenommen werden. Bei den Mädchen kann beobachtet werden, dass sie sich in den *ToM-Aufgaben*, unabhängig von den einzelnen Stereotype Threat-Bedingungen, durchwegs homogen und in Bezug auf die Burschen besser einschätzen. In den Bedingungen *Neutral* und *Aufklärung* kann davon ausgegangen werden, dass diese Unterschiede nicht signifikant kleiner sind als in der Bedingung *Aktivierung*.

Abbildung 12 veranschaulicht die beobachtete Wechselwirkung bei den Aufgaben *Diagramme und Tabellen*. Hier ist anzumerken, dass sich Burschen im Vergleich zu Mädchen in den vier Stereotype Threat-Bedingungen unterschiedlich gut einschätzen.

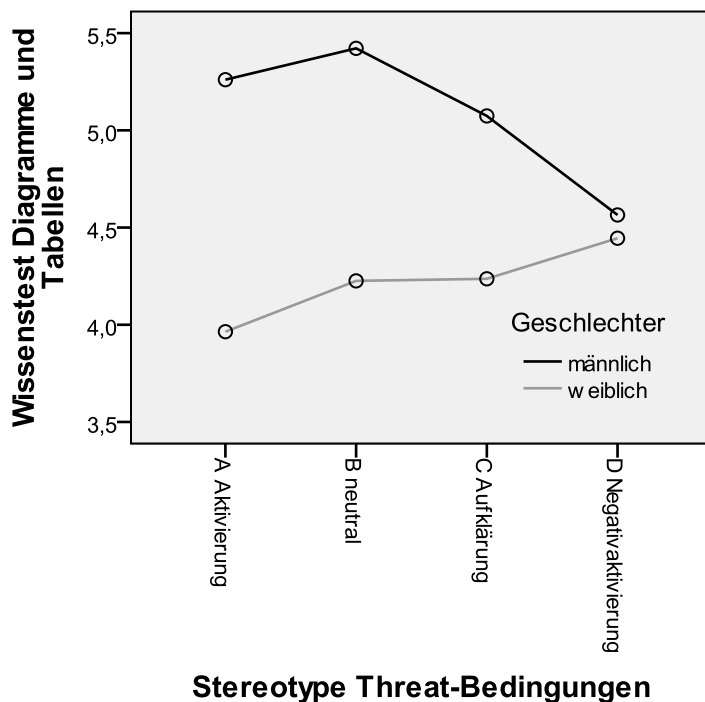


Abbildung 12: Interaktionsdiagramm Diagramme und Tabellen

Zusammenfassend kann man sagen, dass die Überprüfung der Aktivierungswirksamkeit der vier Stereotype Threat-Bedingungen mittels Wissenstest erfolgreich ausgefallen ist und somit die Wirksamkeit der Stereotype Threat-Bedingungen angenommen werden kann. Im Rahmen der Prüfung des Selbstratings zeigte sich, dass die vier Stereotype Threat-Bedingungen nur auf zwei Aufgaben *Diagramme und Tabellen* und *ToM* Auswirkungen hatten. Abschließend ist nochmals zu erwähnen, dass sowohl bei der Fremd- als auch bei der Selbsteinschätzung Frauen und Mädchen eine bessere und ausgeprägtere Fähigkeit in der *ToM* zugeschrieben wird.

8.3 Deskriptive Ergebnisse der Erhebungsinstrumente

Im diesem Abschnitt werden die Ergebnisse der Erhebungsinstrumente deskriptivstatistisch in Bezug auf die Gesamtstichprobe beschrieben. Eingangs werden deskriptive Ergebnisse der EMS-ähnlichen Aufgaben vorgestellt, gefolgt von deskriptivstatistischen Beschreibungen der Ergebnisse des GDT und der

ToM-Aufgaben. Auf die Deskriptivstatistik des LMI und SWE wird an dieser Stelle verzichtet, da diese Aspekte Schwerpunkte der Kolleginnen Gabriele Hangl und Anita Teufl darstellen und in deren Arbeiten ausführlich behandelt werden.

8.3.1 Ergebnisse der EMS-ähnlichen Aufgaben

An dieser Stelle wird auf die Verteilungen der EMS-Scores innerhalb der EMS-Bereiche eingegangen.

Aus Tabelle 51 ist ersichtlich, dass aufgrund des beobachteten Schiefekoeffizienten innerhalb der sechs EMS-Untertests, die Normalverteilung der Daten in der Gesamtstichprobe (N = 1093) angenommen werden kann.

Tabelle 51: statistische Kennwerte der EMS-Untertests gesamt

Statistiken der EMS-Untertests							
		EMS Quant	EMS Schlauch	EMS Natur	EMS Diagramme	EMS Figuren	EMS Fakten
N	Gültig	1093	1093	1093	1093	1093	1093
	Fehlend	0	0	0	0	0	0
M		3.00	10.19	2.81	3.01	4.43	4.24
Md		3.00	10.00	3,00	3.00	4.00	4.00
SD		1.68	3.45	1.69	1.63	2.05	1.77
Schiefe		,485	-,168	,482	,370	-,068	-,111
Standardfehler der Schiefe		,074	,074	,074	,074	,074	,074
Kurtosis		,155	-,283	-,150	-,037	-,638	-,426
Standardfehler der Kurtosis		,148	,148	,148	,148	,148	,148
Minimum		0	0	0	0	0	0
Maximum		10	20	9	9	8	8

8.3.1.1 Quantitative und formale Probleme

Tabelle 52 gibt an, dass die Gesamtstichprobe mit 1093 gültigen Fällen ($M = 3.00$, $SD = 1.67$) in die Berechnungen und Darstellungen mit einbezogen werden kann. Es zeigt sich, dass die getesteten SchülerInnen im Mittel 3 Aufgaben lösen können.

Tabelle 52: Statistik Quantitative und formale Probleme

N	Gültig	1093
	Fehlend	0
M		3,00
Md		3,00
SD		1,67
Schiefe		,485
Minimum		,00
Maximum		10,00

Aus Tabelle 53 lässt sich ablesen, dass 52 der Untersuchten (4,8%) keine einzige Aufgabe dieses Untertests lösen konnten. Der Großteil der TestteilnehmerInnen (266 Personen = 24,3 %) erreichte drei Punkte.

Tabelle 53: Häufigkeiten UT Quantitative und formale Probleme

		Häufigkeit	Prozent	Gültige Prozente
Gültig	,00	52	4,8	4,8
	1,00	155	14,2	14,2
	2,00	238	21,8	21,8
	3,00	266	24,3	24,3
	4,00	185	16,9	16,9
	5,00	111	10,2	10,2
	6,00	55	5,0	5,0
	7,00	22	2,0	2,0
	8,00	7	,6	,6
	9,00	1	,1	,1
	10,00	1	,1	,1
	Gesamt		1093	100,0

Abbildung 13 zeigt eine grafische Darstellung der Verteilung der Testkennwerte des Untertests *Quantitative und formale Probleme* innerhalb der Gesamtstichprobe (N = 1093).

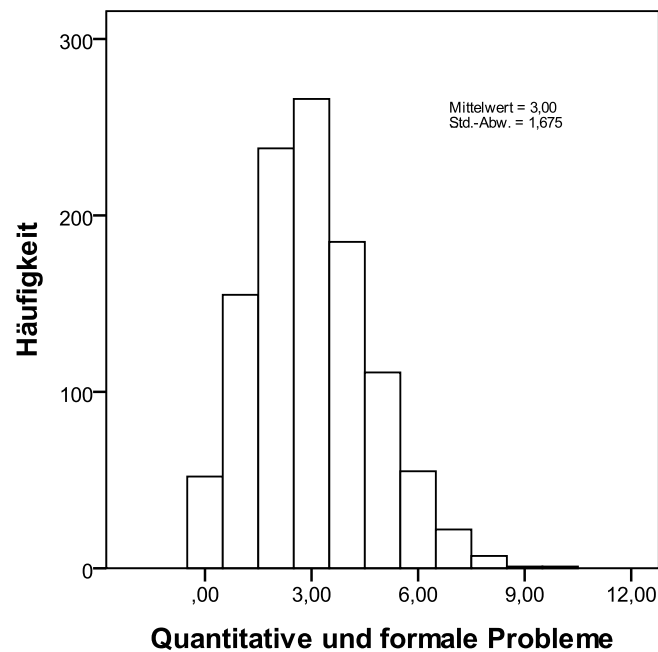


Abbildung 13: Verteilung Quantitative und formale Probleme

8.3.1.2 Schlauchfiguren

Es können 1093 gültige Fälle ($M = 10.19$, $SD = 3.45$) in die Darstellungen des Untertests *Schlauchfiguren* mit einbezogen werden (siehe Tabelle 54).

Tabelle 54: statistische Kennwerte UT Schlauchfiguren

N	Gültig	1093
	Fehlend	0
M		10.19
Md		10.00
SD		3.45
Schiefe		-,168
Minimum		,00
Maximum		20,00

Abbildung 14 veranschaulicht, dass eine Normalverteilung der Kennwerte innerhalb der Gesamtstichprobe angenommen werden kann.

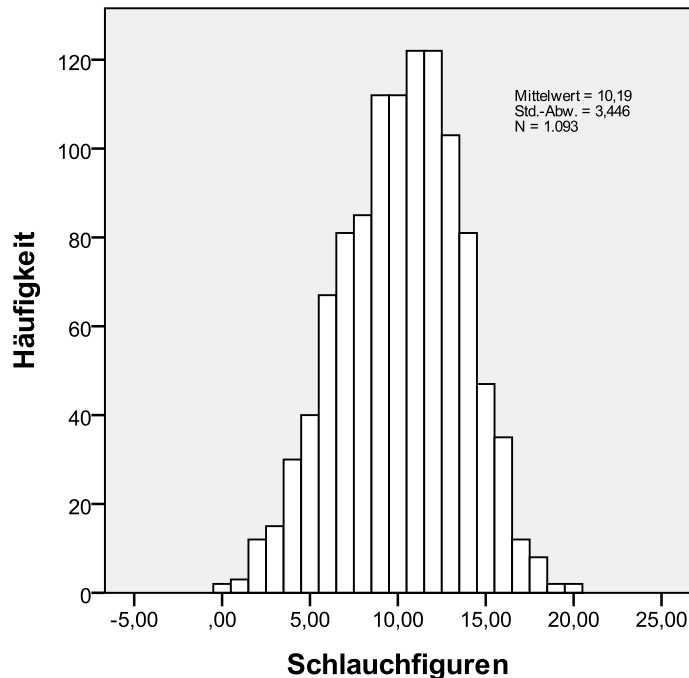


Abbildung 14: Verteilung Schlauchfiguren in der Gesamtstichprobe

In diesem Untertest konnten nur zwei Testpersonen (0,2 %) keine einzige Aufgabe lösen. Des Weiteren kann man aus Tabelle 55 ablesen, dass ebenfalls zwei TestteilnehmerInnen alle Aufgaben bearbeitet und gelöst haben. 122 UntersuchungsteilnehmerInnen (11,2 %) bewältigten zwischen 11 und 12 Aufgaben erfolgreich.

Tabelle 55: Häufigkeiten und Prozentwerte UT Schlauchfiguren

		Häufigkeit	Prozent	Gültige Prozente
Gültig	,00	2	,2	,2
	1,00	3	,3	,3
	2,00	12	1,1	1,1
	3,00	15	1,4	1,4

4,00	30	2,7	2,7
5,00	40	3,7	3,7
6,00	67	6,1	6,1
7,00	81	7,4	7,4
8,00	85	7,8	7,8
9,00	112	10,2	10,2
10,00	112	10,2	10,2
11,00	122	11,2	11,2
12,00	122	11,2	11,2
13,00	103	9,4	9,4
14,00	81	7,4	7,4
15,00	47	4,3	4,3
16,00	35	3,2	3,2
17,00	12	1,1	1,1
18,00	8	,7	,7
19,00	2	,2	,2
20,00	2	,2	,2
Gesamt	1093	100,0	100,0

8.3.1.3 Medizinisch-naturwissenschaftliches Grundverständnis

In den Untertest *Medizinisch-naturwissenschaftliches Grundverständnis* konnten 1093 gültige Fälle ($M = 2.81$, $SD = 1.69$) eingeschlossen werden. Die UntersuchungsteilnehmerInnen können im Mittel 2.81 Aufgaben lösen.

In diesem Untertest können 255 (23,3 %) der getesteten SchülerInnen zwei Aufgaben erfolgreich absolvieren. 70 (6,4 %) SchülerInnen bearbeiten keine der vorgegebenen Aufgaben richtig. Eine einzige Person kann alle Aufgaben lösen (siehe Tabelle 56).

Tabelle 56: Häufigkeiten UT Medizinisch-naturwissenschaftliches Grundverständnis

Medizinisch-naturwissenschaftliches Grundverständnis				
		Häufigkeit	Prozent	Gültige Prozenze
Gültig	,00	70	6,4	6,4
	1,00	188	17,2	17,2
	2,00	255	23,3	23,3
	3,00	231	21,1	21,1
	4,00	172	15,7	15,7
	5,00	98	9,0	9,0
	6,00	52	4,8	4,8
	7,00	23	2,1	2,1
	8,00	3	,3	,3
	9,00	1	,1	,1
	Gesamt		1093	100,0

Abbildung 15 zeigt die Verteilung der getesteten SchülerInnen in diesem Untertest.

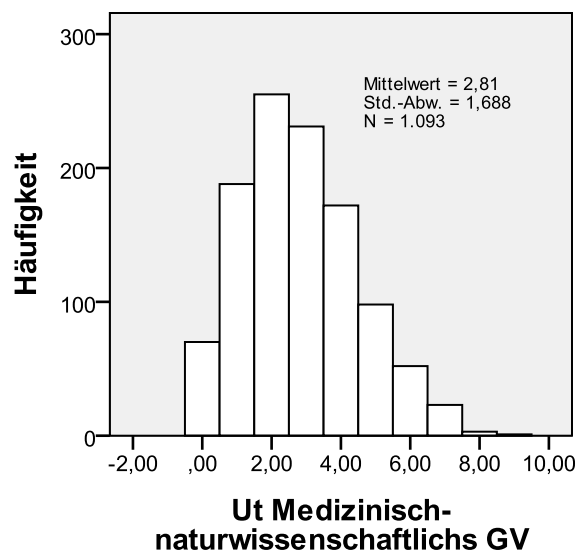


Abbildung 15: Verteilung UT Medizinisch-naturwissenschaftliches GV

8.3.1.4 Diagramme und Tabellen

In diesem Untertest können im Mittel 3.00 Aufgaben ($SD = 1.63$) von den TestteilnehmerInnen (1093 gültige Fälle) gelöst werden. Des Weiteren ist anzumerken, dass die anteilmäßig größte Gruppe mit 257 SchülerInnen (23,5 %) drei Aufgaben richtig bearbeiten. 54 der getesteten SchülerInnen (4,9 %) keine einzige Aufgabe lösen (siehe Tabelle 57).

Tabelle 57: Häufigkeiten UT Diagramme und Tabellen

UT Diagramme und Tabellen				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	,00	54	4,9	4,9
	1,00	143	13,1	13,1
	2,00	243	22,2	22,2
	3,00	257	23,5	23,5
	4,00	207	18,9	18,9
	5,00	104	9,5	9,5
	6,00	63	5,8	5,8
	7,00	18	1,6	1,6
	8,00	1	,1	,1
	9,00	3	,3	,3
	Gesamt	1093	100,0	100,0

In Abbildung 16 wird die Verteilung in diesem Untertest *Diagramme* dargestellt.

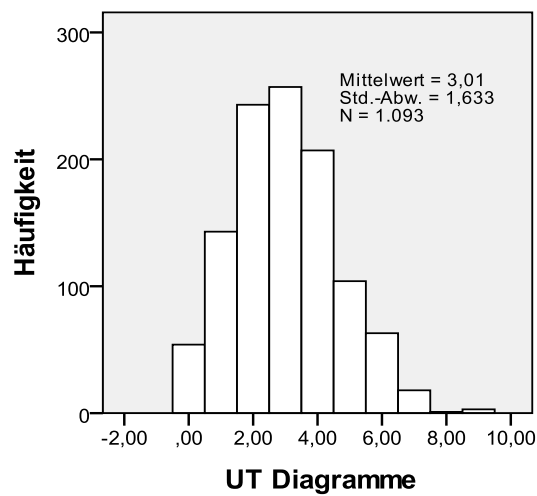


Abbildung 16: Verteilung UT Diagramme und Tabellen

8.3.1.5 Figuren und Fakten lernen

Bei den Untertests *Figuren und Fakten lernen* (siehe Tabelle 58) lösen die UntersuchungsteilnehmerInnen (1093 gültige Fälle) jeweils im Mittel 4.43 ($SD = 2.05$) und 4.24 Aufgaben ($SD = 1.77$).

Tabelle 58: statistische Kennwerte UT Fakten und Figuren

Statistiken Figuren und Fakten			
		Figuren	Fakten
N	Gültig	1093	1093
	Fehlend	0	0
M		4.43	4.24
Md		4:00	4.00
SD		2.05	1.77
Schiefe		-,068	-,111
Standardfehler der Schiefe		,074	,074
Spannweite		8,00	8,00
Minimum		,00	,00
Maximum		8,00	8,00

Anschließend ist aus Tabelle 59 ersichtlich, dass anteilmäßig die meisten SchülerInnen (205 Personen, 18,87%) fünf Aufgaben in diesem Untertest lösen können. 32 der Untersuchten (2,9 %) können keine der Aufgaben erfolgreich bearbeiten.

Tabelle 59: Häufigkeiten UT Figuren lernen

UT Figuren				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	,00	32	2,9	2,9
	1,00	61	5,6	5,6
	2,00	102	9,3	9,3
	3,00	173	15,8	15,8
	4,00	187	17,1	17,1
	5,00	205	18,8	18,8
	6,00	144	13,2	13,2
	7,00	92	8,4	8,4
	8,00	97	8,9	8,9
	Gesamt		1093	100,0

Abbildung 17 veranschaulicht die Verteilung im Untertest *Figuren lernen*.

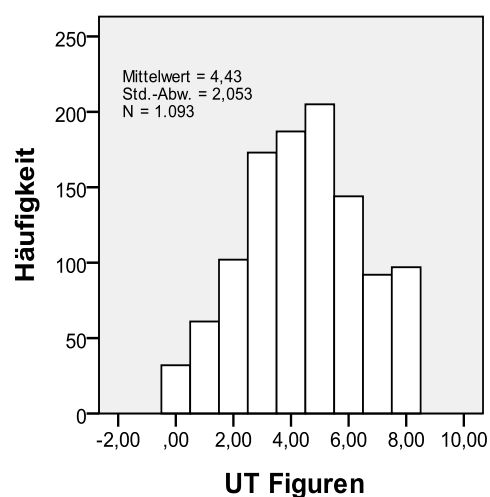


Abbildung 17: Verteilung UT Figuren lernen

Bei dem Untertest *Fakten lernen* zeigen sich ähnliche Ergebnisse (siehe Tabelle 60). Die anteilmäßig größte Gruppe von 232 SchülerInnen (21,2 %) lösen fünf Aufgaben während 21 der Untersuchten (1,9 %) keine Aufgabe richtig bearbeiten.

Tabelle 60: Häufigkeiten UT Fakten lernen

UT Fakten				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	,00	21	1,9	1,9
	1,00	47	4,3	4,3
	2,00	116	10,6	10,6
	3,00	186	17,0	17,0
	4,00	225	20,6	20,6
	5,00	232	21,2	21,2
	6,00	149	13,6	13,6
	7,00	91	8,3	8,3
	8,00	26	2,4	2,4
	Gesamt		1093	100,0

Aus Abbildung 18 ist ersichtlich, dass auch in diesem Untertest von einer Normalverteilung ausgegangen werden kann.

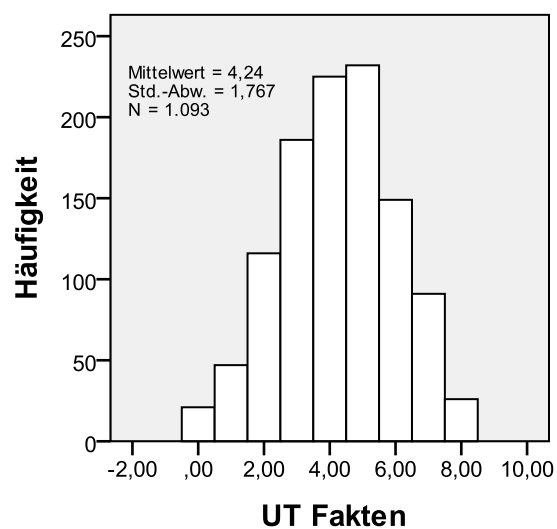


Abbildung 18: Verteilung UT Fakten lernen

8.3.2 Deskriptive Ergebnisse des Game of Dice Task(GDT)

Der GDT konnte 471 Personen ($M = 4.31$, $SD = 9.71$) vorgegeben werden (siehe Tabelle 61).

Tabelle 61: Statistische Kennwerte zum GDT

N	Gültig	471
	Fehlend	622
M		4,31
Md		6,00
SD		9,71
Schiefe		-,52
Minimum		-18
Maximum		18

In Abbildung 19 wird die Verteilung des Netto-Summscore des GDT in der Teilstichprobe ($N = 471$) mittels eines Histogramms veranschaulicht.

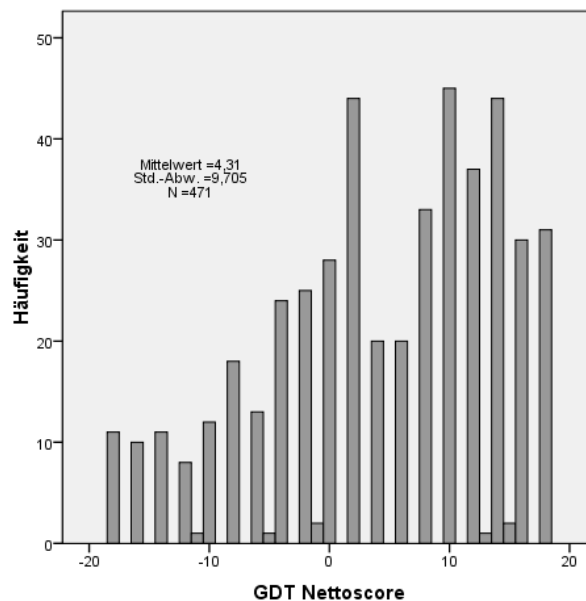


Abbildung 19: Verteilung GDT-Nettoscore

8.3.3 Deskriptivstatistik ToM-Aufgaben

Aus der Gesamtstichprobe (N = 1093) können 1085 gültige Fälle mit einem Mittelwert von 4.77 ($SD = 2.10$) ermittelt werden, die in die nachfolgenden Berechnungen mit einbezogen werden (siehe Tabelle 62).

Tabelle 62: gültige Fälle

N	Gültig	1085
	Fehlend	8
M		4.77
Md		5.00
SD		2.10
Schiefe		-.001
Kurtosis		-.677
Spannweite		9
Minimum		.00
Maximum		9

Unter Berücksichtigung des Schiefekoeffizienten in Tabelle 62 und aus Abbildung 20 ist ersichtlich, dass von einer Normalverteilung der ToM-Aufgaben ausgegangen werden kann.

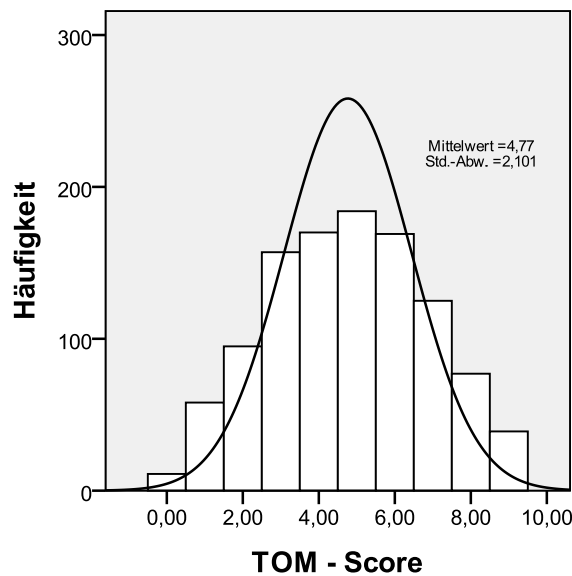


Abbildung 20: Histogramm zum ToM-Score

8.4 Ergebnisdarstellung der Hypothesen

Im folgenden Abschnitt werden die in Kapitel 5 formulierten Hypothesen mittels univariater statistischer Methoden auf ihre Geltung überprüft. Eingangs erfolgt die Prüfung der interessierenden Haupt- und Nebenhypothesen zur ToM durch univariate Varianzanalysen. Im Anschluss daran werden die Ergebnisse der Hypothesenprüfung zu den EMS-Bereichen, sowie die Nebenhypothesen zum GDT dargestellt.

8.4.1 Univariante Ergebnisse der ToM

Wie bereits eingangs erwähnt, erfolgte die Prüfung der Hypothesen mittels univariater ANOVA. Mittels Varianzanalyse wird der Einfluss von einer oder mehreren nominalskalierten unabhängigen Variablen auf eine intervallskalierte abhängige Variable geprüft. Die unabhängigen Variablen werden dabei als Faktoren bezeichnet und die einzelnen Ausprägungen als Faktorstufen. Voraussetzungen für die Varianzanalyse sind die Normalverteilung, die Homogenität der Varianzen sowie Intervallskalenniveau (Backhaus, 2011, S 161).

Es kann von einer Normalverteilung der ToM-Werte in der Gesamtstichprobe ($N = 1085$ gültige Fälle) ausgegangen werden. Der Levene-Test zur Überprüfung der Homogenität der Varianzen fällt für die relevanten Haupthypothesen nicht signifikant aus ($p = .217$), womit die Voraussetzungen für die Berechnung der Varianzanalyse erfüllt sind (siehe Tabelle 63).

Tabelle 63: Levene-Test

Levene-Test auf Gleichheit der Fehlervarianzen ^a			
Abhängige Variable: ToM			
<i>F</i>	df1	df2	Sig.
1,220	23	1061	$p=.217$

Der Levene – Test wird zur Überprüfung der Homogenität von Varianzen im Rahmen von t-Tests und Varianzanalysen eingesetzt. Ein signifikantes Ergebnis in diesem Test bedeutet, dass die Homogenität der Varianzen als verletzt zu betrachten ist und somit eine zentrale Voraussetzung für t-Tests und Varianzanalysen nicht gegeben ist. Fällt das Ergebnis nicht signifikant aus, so kann die Homogenität der Varianzen angenommen werden (Backhaus, 2011).

Geschlecht

Aus Tabelle 64 können die statistischen Kennwerte entnommen werden.

Tabelle 64: Statistische Kennwerte Geschlecht und Stereotype Threat Bedingungen

Deskriptive Statistiken					
Abhängige Variable:ToM					
Geschlecht	Stereotype	Risikoverhalten	<i>M</i>	<i>SD</i>	N
männlich	A Aktivierung	raten	4,51	2.42	35
		ohne Instruktion	4,71	2.07	51
		sicher	3,59	1.89	37
		Gesamt	4,32	2.16	123
	B neutral	raten	4,44	1.83	34
		ohne Instruktion	4,23	2.39	31
		sicher	4,26	2.05	19
		Gesamt	4,32	2.08	84
	C Aufklärung	raten	5,10	2.21	40
		ohne Instruktion	5,28	2,433	29
		sicher	4,79	1,894	42
		Gesamt	5,03	2,147	111
	D Negativaktivierung	raten	4,17	2,320	35
		ohne Instruktion	3,68	2,750	19
		sicher	4,55	2,306	33
		Gesamt	4,21	2,407	87
	Gesamt	raten	4,58	2,211	144
		ohne Instruktion	4,57	2,366	130
		sicher	4,31	2,061	131
		Gesamt	4,49	2,213	405
weiblich	A Aktivierung	raten	4,62	1,927	58

		ohne Instruktion	5,07	1,808	56
		sicher	4,78	1,939	54
		Gesamt	4,82	1,890	168
	B neutral	raten	5,16	1,932	55
		ohne Instruktion	5,27	2,102	56
		sicher	5,25	1,823	61
		Gesamt	5,23	1,941	172
	C Aufklärung	raten	4,66	2,255	64
		ohne Instruktion	4,66	2,297	62
		sicher	5,55	1,837	42
		Gesamt	4,88	2,196	168
	D Negativaktivierung	raten	4,45	1,653	51
		ohne Instruktion	5,06	2,071	62
		sicher	4,86	2,193	59
		Gesamt	4,81	2,006	172
	Gesamt	raten	4,72	1,975	228
ohne Instruktion		5,01	2,081	236	
sicher		5,08	1,970	216	
Gesamt		4,94	2,014	680	
Gesamt	A Aktivierung	raten	4,58	2,113	93
		ohne Instruktion	4,90	1,937	107
		sicher	4,30	1,997	91
		Gesamt	4,61	2,022	291
	B neutral	raten	4,89	1,915	89
		ohne Instruktion	4,90	2,252	87
		sicher	5,01	1,913	80
		Gesamt	4,93	2,028	256
	C Aufklärung	raten	4,83	2,236	104
		ohne Instruktion	4,86	2,346	91
		sicher	5,17	1,894	84
		Gesamt	4,94	2,174	279
	D Negativaktivierung	raten	4,34	1,944	86
		ohne Instruktion	4,74	2,306	81
		sicher	4,75	2,227	92
		Gesamt	4,61	2,163	259
	Gesamt	raten	4,67	2,068	372
		ohne Instruktion	4,85	2,193	366
		sicher	4,79	2,037	347
		Gesamt	4,77	2,101	1085

Die ANOVA zur Überprüfung der $H_1(1)$ ergibt mit der entsprechenden Prüfgröße $F(1,1061) = 14.316$, $p = .0001$ ein signifikantes Ergebnis (siehe Tabelle 65). Daraus resultiert, dass ein signifikanter Unterschied in der ToM in Abhängigkeit vom Geschlecht angenommen werden kann. Die getesteten Burschen und Mädchen unterscheiden sich signifikant in der ToM, operationalisiert durch die ToM-Aufgaben. Die weiblichen Schüler weisen höhere Werte in der ToM ($M = 4.94$, $SD = 2.01$, 680 gültige Fälle, siehe Tabelle xy) im Vergleich zu den männlichen Schülern auf ($M = 4.49$, $SD = 2.21$, 405 gültige Fälle).

Tabelle 65: Ergebnisse der ANOVA

Tests der Zwischensubjekteffekte						
Abhängige Variable: ToM						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	200,400 ^a	23	8,713	2,017	.003	,042
Konstanter Term	21218,120	1	21218,120	4911,091	.000	,822
Geschlecht	61,851	1	61,851	14,316	$p \leq .0001$,013
Stereotype	43,795	3	14,598	3,379	$p = .018$,009
Risikoverhalten	1,848	2	,924	,214	$p = .807$,000
Geschlecht* Stereotype	34,088	3	11,363	2,630	$p = .049$,007
Geschlecht* Risikoverhalten	17,334	2	8,667	2,006	$p = .135$,004
Stereotype * Risikoverhalten	31,799	6	5,300	1,227	$p = .290$,007
Geschlecht* Stereotype * Risikoverhalten	32,722	6	5,454	1,262	$p = .272$,007
Fehler	4583,997	1061	4,320			
Gesamt	29467,000	1085				
Korrigierte Gesamtvariation	4784,396	1084				
a. R-Quadrat = ,042 (korrigiertes R-Quadrat = ,021)						

Risikoverhalten

Die Prüfung der $H_1(2)$ ergibt mit $F(2, 1061) = .214, p = .807$ kein signifikantes Ergebnis (siehe Tabelle 65). Es kann somit kein Unterschied in der ToM in Abhängigkeit von den Ausprägungen des Risikoverhaltens beobachtet werden.

Stereotype Threat

Die ANOVA zur Prüfung der $H_1(3)$ ergibt ein signifikantes Ergebnis (siehe Tabelle 65). Es kann, gemäß der Berechnung der entsprechenden Prüfgröße $F(3, 1061) = 3.379, p = .018$ ein signifikanter Unterschied in der ToM in Abhängigkeit von den vier Stereotype Threat-Bedingungen angenommen werden. Die Berechnung von Post-hoc-Tests nach Bonferroni zeigt jedoch für keinen paarweisen Vergleich signifikante Unterschiede an, was durch die geringere Teststärke von Post-hoc-Tests erklärt werden kann.

Geschlecht und Stereotype Threat

Die Prüfung der $H_1(4)$ ergibt ein signifikantes Ergebnis (siehe Tabelle 65). Es kann eine signifikante Wechselwirkung in den ToM-Aufgaben in Abhängigkeit vom Geschlecht und den Stereotype Threat-Bedingungen mit der Prüfgröße $F(3, 1061) = 2.630, p = .049$ beobachtet werden (siehe Abbildung 21).

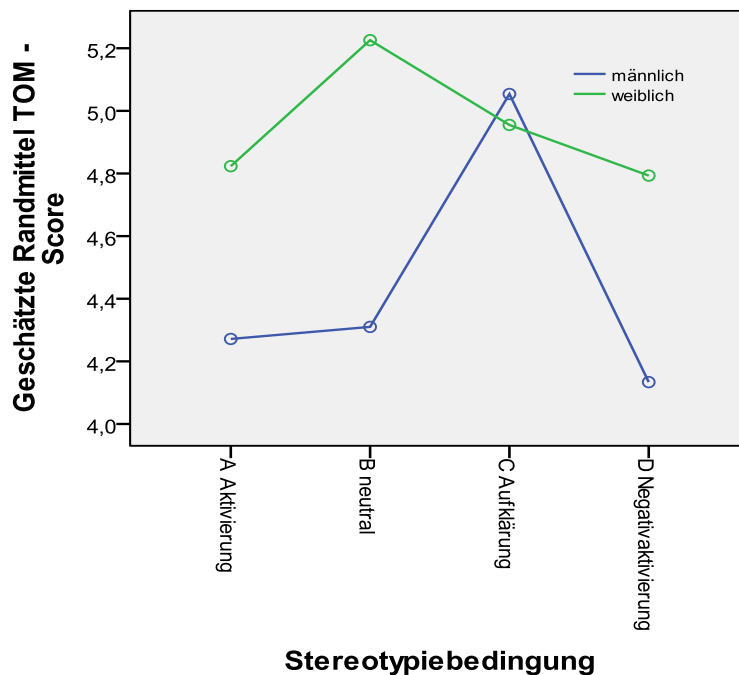


Abbildung 21: Interaktionsdiagramm der WW Stereotype Threat und Geschlecht

Aus dieser Abbildung lässt sich entnehmen, dass die Differenzen in ToM zwischen den Geschlechtern in Abhängigkeit von den Stereotype Threat-Bedingungen signifikant unterschiedlich sind. Burschen weisen in den Stereotype Threat-Bedingungen *Aktivierung*, *Neutrale Bedingung* und *Negativ-Aktivierung* niedrigere Werte in den ToM-Aufgaben auf als Mädchen. Es zeigt sich jedoch, dass diese Niveauunterschiede verschwinden, wenn Burschen in der *Aufklärungsbedingung* über das Phänomen Stereotype Threat informiert und aufgeklärt werden. Grundsätzlich ist anzumerken, dass die Gruppe der Mädchen in allen Bedingungen bis auf die *Aufklärungsbedingung* höhere Werte in den ToM-Aufgaben erzielt als die Gruppe der Burschen. Die höchsten Werte weisen die weiblichen Untersuchungsteilnehmer in der *Neutralen Bedingung* auf. Interessanterweise erreichen Mädchen in der *Aktivierungsbedingung* und in der *Negativ-Aktivierungsbedingung* annähernd gleich hohe Werte.

Geschlecht und Risikoverhalten

Die Überprüfung der H_1 (5) weist mit $F(2, 1061) = 2.006$, $p = .135$ ein nicht signifikantes Ergebnis auf (siehe Tabelle 65). Daraus resultiert die Annahme der H_0 , was bedeutet, dass keine signifikante Interaktion in der ToM in Abhängigkeit vom Risikoverhalten angenommen werden kann.

Stereotype Threat und Risikoverhalten

Die ANOVA zur Klärung der H_1 (6) zeigt mit $F(6, 1061) = 1.227$, $p = .290$ ein nicht signifikantes Ergebnis (siehe Tabelle 65). Daraus folgt, dass keine signifikante Interaktion in der ToM in Abhängigkeit von den Stereotype Threat-Bedingungen und vom Risikoverhalten beobachtet werden kann.

Geschlecht, Stereotype Threat und Risikoverhalten

Die dreifaktorielle ANOVA zur Prüfung der H_1 (7) weist mit $F(6, 1061) = 1.262$, $p = .272$ ein nicht signifikantes Ergebnis auf, womit die H_0 beizubehalten ist (siehe Tabelle 65). Es kann somit keine signifikante Interaktion höherer Ordnung in der ToM in Abhängigkeit vom Geschlecht, den Stereotype Threat-Bedingungen und dem Risikoverhalten angenommen werden.

8.2.2 Zusammenhang zwischen ToM und den EMS-ähnlichen Aufgaben

In diesem Abschnitt erfolgt die Darstellung der Prüfung des Hypothesenblocks $H_1(8-14)$ mittels Produkt-Momentkorrelation nach Pearson.

Mit diesem Verfahren wird geprüft, ob ein linearer Zusammenhang zwischen zwei intervallskalierten Variablen besteht. Aus dem Korrelationskoeffizienten kann die Stärke und Richtung dieses Zusammenhangs abgeleitet werden, wobei Korrelationen von 0.1 bis 0.29 für geringe, Korrelationen von 0.3 bis 0.49 für mittlere und Korrelationen ≥ 0.5 für hohe Zusammenhänge stehen (Bortz & Döring, 2006, S. 606). Die Voraussetzungen für die Produkt-Momentkorrelation

nach Pearson sind Intervallskalenniveau und Normalverteilung (Bortz & Döring, 2006, S.508).

In der folgenden Tabelle 66 werden die Koeffizienten der Produktmomentkorrelation von ToM mit den sechs EMS-Untertests dargestellt.

Tabelle 66: Interkorrelationsmatrix der Koeffizienten der Produkt-Momentkorrelation ToM mit EMS-ähnlichen Aufgaben(n = 1085)

		EMS Quant	EMS Schlauch	EMS Natur	EMS Diagramm e	EMS Figuren	EMS Fakten
ToM	Korrelation nach Pearson	.070*	.289**	.150**	.087**	.203**	.204**
	Signifikanz (2- seitig)	$p = ,022$	$p \leq ,0001$	$p \leq ,0001$	$p = ,004$	$p \leq ,0001$	$p \leq ,0001$

Es können jeweils sehr geringe bis niedrige, positive signifikante Zusammenhänge zwischen ToM und den EMS-ähnlichen Aufgaben beobachtet werden. Dies bedeutet, dass mit höheren ToM-Leistungen im Allgemeinen auch höhere EMS-Werte einhergehen.

Mittels multipler linearer Regression werden zur Vorhersage des ToM-Kriteriums (abhängige Variable) die sechs EMS-Untertests als unabhängige Prädiktorvariablen schrittweise in ein Modell aufgenommen. Die Voraussetzungen zur Modellprüfung, die Normalverteilung der standardisierten Residuen sowie keine Multikollinearität der unabhängigen Variablen können angenommen werden.

Mittels multipler linearer Regression erfolgt die Überprüfung des Einflusses von mehreren intervallskalierten (oder nominalskalierten) unabhängigen Variablen (Prädiktoren) auf eine intervallskalierte abhängige Variable (Kriteriumsvariable). Das Ausmaß der Modellpassung wird durch den erklärten Varianzanteil R^2 angegeben. Zur Erfassung der Bedeutung und der Gewichtung der einzelnen Variablen werden die Regressionskoeffizienten standardisiert dargestellt. Die Voraussetzungen lineare Zusammenhänge zwischen abhängiger Variable (AV) und ihren Prädiktoren (UVs), keine Multikollinearität, worunter keine lineare

Abhängigkeit der Prädiktoren untereinander verstanden wird, annähernd normalverteilte Residuen, keine Korrelation zwischen den unabhängigen Variablen und Residuen sowie Homoskedastizität, was bedeutet, dass die Streuung der Residuen in verschiedenen Bereichen der unabhängigen Variablen nicht unterschiedlich ist (Bortz & Döring, 2006).

Die globale Modellzusammenfassung fällt mit $F(4, 1080) = 39.328, p \leq .0001$ signifikant aus, es können vier Prädiktoren mit Erklärungswert identifiziert werden.

Die Tabelle 67 zeigt die Koeffizienten und Parameter der in das Modell aufgenommenen Variablen.

Tabelle 67: in das Modell aufgenommene Variablen

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	t	Sig.
		Regr.koeffizient B	Standardfehler SE	β		
4	(Konstante)	1.893	.238		7.958	$p \leq .0001$
	EMS Schlauch	.141	.018	.230	7.746	$p \leq .0001$
	EMS Fakten	.145	.036	.120	4.010	$p \leq .0001$
	EMS Figuren	.130	.031	.126	4.232	$p \leq .0001$
	EMS Natur	.087	.037	.070	2.373	$p = .018$

Der erklärte Varianzanteil an der abhängigen Variable ToM beträgt $R^2 = 12,7\%$. Unter Berücksichtigung der standardisierten β – Gewichte ist erkennbar, dass EMS Schlauchfiguren den relativ bedeutendsten Anteil für die Prognosefähigkeit aufweisen. Die EMS-Variablen *Quantitative und formale Probleme* und *Diagramme und Tabellen* weisen keinen signifikanten Erklärungswert auf und werden aus dem Modell ausgeschlossen (siehe Tabelle 68).

Tabelle 68: aus dem Modell ausgeschlossene Variablen

Modell		Beta In	t	Sig.
4	EMS Quant	-.010 ^d	-.325	$p = .745$
	EMS Diagramme	.003 ^d	.102	$p = .919$

8.2.3 Nebenhypothesen zur ToM

Im Folgenden wird die Prüfung der interessierenden Nebenhypothesen zur ToM, mittels t – Test und zweifaktorieller Varianzanalyse, dargestellt.

Mittels t-Test wird eine Unterschiedshypothese durch den Vergleich von Mittelwerten der beiden Gruppen geprüft. Es wird die Signifikanz der Mittelwertdifferenz zwischen zwei Gruppen untersucht. Für die Anwendung dieses Verfahrens müssen die Voraussetzungen der Normalverteilung, des Intervallskalenniveaus, die Unabhängigkeit der Testwerte und die Homogenität der Varianzen erfüllt sein (Bortz & Döring, 2006).

Wunsch, Medizin zu studieren

Die Prüfung der H_1 (8) mittels t-Test zeigt auf Grundlage jener 1060 gültigen Fälle, die einen Studienwunsch (ja/nein) für ein zukünftiges Medizinstudium angegeben haben, dass die Prüfung der Varianzhomogenität mit $p = .011$ signifikant ausfällt und somit die Homogenität der Varianzen als verletzt zu betrachten ist. Aus Tabelle 69 können die statistischen Kennwerte entnommen werden.

Tabelle 69: Statistische Kennwerte

Gruppenstatistiken					
	Medizin	N	M	SD	Standardfehler
ToM	nein	873	4.78	2.13	.072
	ja	187	4.79	1.90	.139

Zur Klärung der Hypothese wird infolgedessen ein Welch-Test für unabhängige Stichproben angewendet. Der Welch – Test ist ein t- Test für heterogene Varianzen (Bortz & Döring, 2006, S 496). Die Berechnung der Prüfgröße ergibt mit $t(295.26) = - 0.043$, $p = . 965$ keinen signifikanten Unterschied in der ToM in Abhängigkeit von dem Wunsch, Medizin zu studieren (siehe Tabelle 70). Daraus lässt sich ableiten, dass die geplante Studienwahl keinen Einfluss auf das Bewältigen der ToM-Aufgaben hat.

Tabelle 70 :Welch-Test

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Sig.	T	df	Sig. (2-seitig)	Mittlere Differenz z	Standardfehler der Differenz z	95% Konfidenzintervall der Differenz	
									Untere	Obere
ToM	Varianzen sind gleich	6,424	$p=.011$	-,040	1058	$p=.968$	-,007	,168	-,337	,323
	Varianzen sind nicht gleich			-,043	295,257	$p=.965$	-,007	,156	-,314	,301

Geschlecht und Wunsch nach Medizinstudium

Mittels zweifaktorieller Varianzanalyse wird darüber hinaus geprüft, ob Geschlecht und Wunsch, nach Medizinstudium einen Effekt auf ToM zeigen. In Tabelle 71 werden die statistischen Kennwerte dargestellt.

Tabelle 71 Statistische Kennwerte Geschlecht und Wunsch nach Medizinstudium

Deskriptive Statistiken				
Abhängige Variable: ToM				
Geschlecht	Medizin	<i>M</i>	<i>SD</i>	N
männlich	nein	4.43	2.20	330
	ja	5.00	2.06	62
	Gesamt	4.52	2.19	392
weiblich	nein	5.00	2.05	543
	ja	4.69	1.81	125
	Gesamt	4.94	2.01	668
Gesamt	nein	4.78	2.13	873
	ja	4.79	1.90	187
	Gesamt	4.79	2.09	1060

Die Prüfung der Homogenität der Varianzen mittels Levene - Test fällt mit $p = .022$ signifikant aus. Die Varianzanalyse ist gegenüber leichten Verletzungen der Homogenität robust und daher können die Ergebnisse interpretiert werden. Die Prüfung der letzten Nebenhypothese $H_1(10)$ zur ToM ergibt mit entsprechender Prüfgröße $F(1,1056) = 6.23$, $p = .013$ ein signifikantes Ergebnis (siehe Tabelle 72).

Tabelle 72. ANOVA Geschlecht und Wunsch nach Medizinstudium

Tests der Zwischensubjekteffekte						
Abhängige Variable: ToM						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	<i>F</i>	<i>Sig.</i>	Partielles Eta-Quadrat
Korrigiertes Modell	70,659 ^a	3	23,553	5,471	$p = .001$,015
Konstanter Term	12603,361	1	12603,361	2927,836	$p \leq .000$,735
Geschlecht	2,290	1	2,290	,532	$p = .466$,001
Medizin	2,290	1	2,290	,532	$p = .466$,001
Geschlecht * Medizin	26,806	1	26,806	6,227	$p = .013$,006

Fehler	4545,729	1056	4,305			
Gesamt	28895,000	1060				
Korrigierte Gesamtvariation	4616,388	1059				
a. R-Quadrat = ,015 (korrigiertes R-Quadrat = ,013)						

Daraus folgt, dass eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und dem Wunsch, Medizin zu studieren angenommen werden kann (siehe Abbildung 22). Der Unterschied im ToM-Score in den Geschlechtern ist in Abhängigkeit vom Wunsch, Medizin zu studieren, signifikant unterschiedlich.

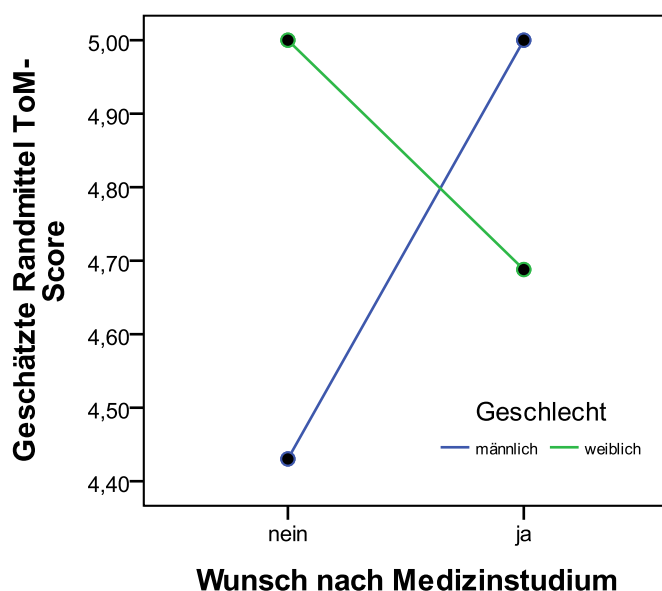


Abbildung 22: Interaktionsdiagramm für ToM in Abhängigkeit vom Geschlecht und Wunsch, nach Medizinstudium

Um eine Aussage darüber treffen zu können, wo genau diese Unterschiede liegen, werden t-Tests für unabhängige Stichproben, jeweils für die Gruppe *ohne Wunsch Medizin zu studieren* und für die Gruppe *mit dem Wunsch nach einem Medizinstudium*, durchgeführt. In die Berechnungen werden insgesamt $N = 873$ gültige Fälle mit einbezogen (siehe Tabelle 73).

Tabelle 73: Gruppenstatistiken kein Wunsch nach Medizinstudium

Gruppenstatistiken				
	Geschlecht	N	M	SD
ToM	männlich	330	4.43	2.20
	weiblich	543	5.00	2.05

Die Homogenität der Varianzen kann angenommen werden, da die Prüfung mittels Levene-Test mit $p = .056$ ein nicht signifikantes Ergebnis aufweist. Die Berechnung der entsprechenden Prüfgröße $t(871) = -3.87$, $p \leq .0001$ ergibt ein signifikantes Ergebnis (siehe Tabelle 74). Es kann ein signifikanter Unterschied im ToM-Score innerhalb der Gruppe *ohne Wunsch nach einem Medizinstudium* beobachtet werden. Innerhalb dieser Gruppe zeigt sich, dass Mädchen im ToM-Score höhere Werte als Burschen erzielen. Eine anschließende Effektstärkenberechnung ergibt mit $d = 0.33$ einen kleinen Effekt. Es kann somit nur ein kleiner Effekt, dass Mädchen in der Gruppe *ohne Wunsch nach einem Medizinstudium* besser in den ToM-Aufgaben abschneiden als ihre männlichen Schulkollegen, angenommen werden.

Tabelle 74: t-Test: Gruppe ohne Wunsch nach Medizinstudium

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Sig.	T	df	Sig. (2-seitig)	Mittlere Differenz z	Standardfehler der Differenz z	95% Konfidenzintervall der Differenz	
									Untere	Obere
ToM	Varianzen sind gleich	3,669	.056	-3,867	871	$p \leq .0001$	-,56970	,14734	-,85887	-,28052
	Varianzen sind nicht gleich			-3,800	656,223	$p \leq .0001$	-,56970	,14994	-,86411	-,27529

Für die Berechnung des zweiten t-Tests für die Gruppe *Wunsch nach Medizinstudium* werden $N = 185$ gültige Fälle eingeschlossen (siehe Tabelle 75). Auch hier kann die Homogenität der Varianzen mit dem nicht signifikanten Ergebnis von $p = .194$ des Levene – Tests angenommen werden.

Tabelle 75: Gruppenstatistiken Wunsch nach Medizinstudium

Gruppenstatistiken				
	Geschlecht	N	M	SD
FB gesamt	männlich	62	5.00	2.06
	weiblich	125	4.69	1.81

Das Ergebnis des t-Tests fällt mit der entsprechenden Prüfgröße $t(185) = 1.06$, $p = .291$ nicht signifikant aus (siehe Tabelle 76). Daraus folgt, dass kein signifikanter Unterschied der Geschlechter im ToM-Score innerhalb der Gruppe *Wunsch nach einem Medizinstudium* beobachtet werden kann.

Tabelle 76: t-Test: Gruppe Wunsch nach einem Medizinstudium

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Sig.	T	df	Sig. (2-seitig)	Mittlere Differenz z	Standardfehler der Differenz z	95% Konfidenzintervall der Differenz	
									Untere	Obere
ToM	Varianzen sind gleich	1,700	$p=.194$	1,060	185	$p=.291$,31200	,29447	-,26895	,89295
	Varianzen sind nicht gleich			1,015	109,030	$p=.312$,31200	,30735	-,29716	,92116

8.2.4 Nebenhypothese zum GDT

Die Prüfung der Nebenhypothesen H_1 (17,18,19) erfolgt mittels zweifaktorieller ANOVA. In die Berechnungen können 471 gültige Fälle mit einbezogen werden. Im Zuge der explorativen Datenanalyse, gesplittet nach Gruppe, zeigt sich, dass die Normalverteilung der Daten je Gruppe angenommen werden kann. Der Levene – Test ergibt mit $p = .922$ ein nicht signifikantes Ergebnis, womit die Homogenität der Varianzen gegeben ist. Die Voraussetzungen für die Berechnung der Varianzanalyse sind damit gegeben. In Tabelle 77 werden die Mittelwerte der drei Risikobedingungen dargestellt. Es zeigt sich, dass die Rate-Bedingung den niedrigsten Mittelwert aufweist, was auf risikoreiches Verhalten hindeuten könnte, da niedrige Werte im GDT für riskantes Entscheidungsverhalten stehen.

Tabelle 77: Statistik Risikobedingungen in N = 471

Deskriptive Statistiken				
Abhängige Variable: ns_GDT				
	Risikoverhalten	<i>M</i>	<i>SD</i>	N
Gesamt	raten	2.72	9.67	157
	ohne Instruktion	5.28	9.31	163
	sicher	4.91	10.01	151
	Gesamt	4.31	9.71	471

Die Prüfung mittels ANOVA ergibt für den Faktor Geschlecht, mit $F(1, 465) = 3.62$, $p = .0580$ ein knapp nicht signifikantes Ergebnis ($H_1(18)$), was bedeutet, dass kein signifikanter Unterschied im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit vom Geschlecht festgestellt werden kann (siehe Tabelle 78). Es kann jedoch eine tendenzielle Signifikanz bezüglich des Geschlechts angenommen werden. Die anschließende Effektstärkenberechnung ergibt $d = 0,18$ einen geringen Effekt.

Tabelle 78: ANOVA ns.GDT(Risikofreude), Geschlecht und Risikoverhalten

Tests der Zwischensubjekteffekte						
Abhängige Variable:ns_GDT						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	975,765 ^a	5	195,153	2,096	$p=.065$,022
Konstanter Term	7567,904	1	7567,904	81,290	$p\leq.000$,149
Geschlecht	336,981	1	336,981	3,620	$p=.058$,008
Risikoverhalten	633,373	2	316,687	3,402	$p=.034$,014
Geschlecht * Risikoverhalten	36,736	2	18,368	,197	$p=.821$,001
Fehler	43290,596	465	93,098			
Gesamt	53007,000	471				
Korrigierte Gesamtvariation	44266,361	470				

a. R-Quadrat = ,022 (korrigiertes R-Quadrat = ,012)

Die Überprüfung der $H_1(17)$ ergibt mit entsprechender Prüfgröße $F(2, 465) = 3.40$, $p = .034$ ein signifikantes Ergebnis. Es kann ein signifikanter Unterschied im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit von der Risikoinstruktion angenommen werden. Des Weiteren zeigt die Prüfung der $H_1(19)$, dass keine signifikante Wechselwirkung mit $F(2, 465) = 0.197$, $p = .821$ im Risikoverhalten, operationalisiert durch den GDT, in Abhängigkeit vom Geschlecht und der Risikoinstruktion angenommen werden kann (siehe Abbildung 23).

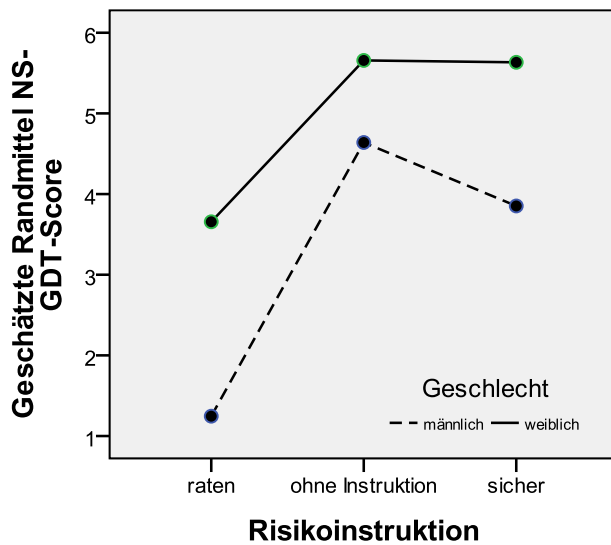


Abbildung 23: NS GDT- Score Risikofreude (NS GDT-Score) in Abhängigkeit vom Risikoverhalten und Geschlecht

Die Betrachtung der Abbildung 23 zeigt jedoch, dass Männer tendenziell risikobereiter sind. Mittels a priori-Kontraste, *simple first – simple last*, werden, zur Identifizierung der Unterschiede in der Risikobereitschaft, in Abhängigkeit von der Risikoinstruktion, die Stufen paarweise miteinander verglichen, da die Erwartung besteht, dass sich die *Rate*-Bedingung von den anderen beiden Bedingungen unterscheidet. A-priori- Kontraste setzten die Formulierung gezielter Einzelvergleichshypothesen vor der Untersuchung voraus (Bortz & Döring, S530, 2006).

Es zeigt sich für den Vergleich der Bedingungen *Raten* und *Sicher* ein signifikantes Ergebnis ($p = .042$), was darauf hinweist, dass die UntersuchungsteilnehmerInnen in dieser Bedingung riskanter entscheiden. Bezüglich der Bedingungen *Neutral* und *Aufklärung* kann kein signifikanter Unterschied ($p = .717$) festgestellt werden. Es kann jedoch ein signifikanter Unterschied ($p = .16$) zwischen den Bedingungen *Raten* und *ohne Instruktion* beobachtet werden. Die Vermutung, dass sich die *Rate*-Bedingung signifikant von den anderen beiden Bedingungen unterscheidet, konnte bestätigt werden. Es zeigt sich, dass die TestteilnehmerInnen in dieser Bedingung riskanter entscheiden als

in den anderen zwei Bedingungen. Die Wirksamkeit dieser Risikoinstruktion kann somit angenommen werden.

Kapitel 9 Diskussion:

In diesem Abschnitt werden die Ergebnisse zusammenfassend und interpretativ dargestellt. In weiterer Folge werden Verbindungen und Zusammenhänge zu den im Theorieteil angeführten relevanten und interessierenden Aspekten unter Einbezug der vorhandenen Literatur hergestellt und diskutiert. Im Anschluss daran erfolgt eine kritische Auseinandersetzung in Bezug auf die Vorgehensweise und Durchführung der Untersuchung. Den Abschluss bildet ein Ausblick über mögliche Richtungen, in die weitere Untersuchungen in Bezug auf die in dieser Arbeit dargestellte Thematik gehen könnten.

Das Ziel dieser empirischen Arbeit war die Untersuchung der ToM und die Bearbeitung EMS-ähnlicher Aufgaben in Abhängigkeit des Phänomens Stereotype Threat sowie Risikoverhalten und Geschlecht. Es sollten potenzielle Zusammenhänge und wechselseitige Einflüsse dieser Konstrukte unter Berücksichtigung geschlechtsspezifischer Unterschiede erforscht und beleuchtet werden.

Die Thematik der geschlechtsspezifischen Unterschiede, sowohl in kognitiven als auch in sozial-kognitiven Bereichen, ist in der Literatur ein vielfach diskutiertes Feld. Tatsache ist, dass es diese fähigkeitsbezogenen Unterschiede zwischen den Geschlechtern gibt und Frauen und Männer über unterschiedliche Stärken und Schwächen, Talente und Begabungen verfügen. Das Vorhandensein dieser Unterschiede sollte jedoch für keines der Geschlechter Nachteile und oder negative Bewertungen nach sich ziehen und weder Frauen noch Männer in ihrem persönlichen und beruflichen Werdegang hemmen. Baron-Cohen (2002;2003; Baron-Cohen & Hammer, 1997) veranschaulicht und diskutiert die Thematik der geschlechtsspezifischen Unterschiede im Rahmen der *Extreme male brain-Theorie* und der *Empathizing-Sytsemizing-Theorie*, in dem er diese Unterschiede unter anderem auf die unterschiedliche Entwicklung und Strukturierung der weiblichen und männlichen Gehirne zurückführt. Demnach ist der Bereich des Einfühlens und der Empathiefähigkeit bei Frauen stärker ausgeprägt als bei Männern, bei denen die Fähigkeit und der Bereich des Systematisierens im Vordergrund stehen. Baron-Cohen (2002) kam im Zuge mehrere Untersuchungen

zu dem Schluss, dass Männer und Frauen die Welt unterschiedlich wahrnehmen und begreifen und sich diese Tatsache auch in ihren Leistungen und Fähigkeiten in kognitiven und sozial-kognitiven Bereichen manifestiert. Frauen erklären sich die Welt durch Empathie und mit Hilfe von Gefühlen, während Männer in Systemen denken und sich mit der Funktionalität von Objekten beschäftigen. Geschlechtsspezifische Unterschiede bestehen somit unweigerlich, sollten aber fernab von negativen Bewertungen und den daraus resultierenden Nachteilen betrachtet werden. Die Realität zeichnet jedoch ein anderes Bild. Ein Beispiel dafür liefert der viel diskutierte EMS-Test, in dem, wie Mallinger et al. (2009), sowie Spiel et al. (2007) im Rahmen einer Evaluierung feststellten, Frauen und Mädchen in sieben (*Quantitative und formale Probleme, Schlauchfiguren, Textverständnis, Planen und Organisieren, Konzentriertes und sorgfältiges Arbeiten, Medizinisch-naturwissenschaftliches Verständnis und Diagramme und Tabellen*) von zehn Unterstests systematisch schlechter abschneiden als Männer und Burschen. Diese Untertests erfassen zum Großteil räumlich-visuelle und mathematische Fähigkeiten die, wie bereits vielfach aus der Literatur bekannt ist (Halpern & LaMay, 2000; Halpern, 2004; Bosco et al., 2002, Maccoby & Jacklin, 1974), zu den männlichen Domänen zählen. Die Begabungen der Frauen und Mädchen liegen, wie zum Beispiel Studien von Walker (2005), Baron-Cohen und Wheelwright (2004), sowie Lawton und Hatcher (2005), Lowe et al. (2003) und Huang (1993) bestätigen konnten, im Gedächtnis, in der verbalen Fähigkeit, in der sozialen Kompetenz und in der Empathiefähigkeit. Es kann somit eine klare Benachteiligung der Frauen im Rahmen dieses Aufnahmeverfahrens. In diesem Zusammenhang werden Fragen nach den Gründen und Ursachen, sowie möglichen Einflussfaktoren für diese Unterschiede aufgeworfen.

Im Rahmen dieser größer angelegten Untersuchung sollten mögliche Ursachen und Erklärungen, Einflussfaktoren sowie modulierende Variablen für die Unterschiede im Bewältigen von EMS-ähnlichen und sozial-kognitiven Aufgaben (ToM-Aufgaben) eruiert werden. Die Ergebnisse bezüglich geschlechtsspezifischer Unterschiede in ausgewählten Untertests werden in den Arbeiten von den Kolleginnen Helene Lagger und Silvia Hameseder ausführlich dargestellt und sind dort nachzulesen. Der Einfluss von Variablen wie der

Leistungsmotivation und allgemeinen Selbstwirksamkeitserwartung werden im Rahmen der Arbeiten von Gabriele Hangl und Anita Teufl behandelt worden.

Der Schwerpunkt dieser Arbeit liegt, wie bereits erwähnt, in der Untersuchung der ToM unter Berücksichtigung geschlechtsspezifischer Unterschiede und der Konstrukte Stereotype Threat und Risikoverhalten.

In der Literatur belegen mehrere Studien (Bosacki et al., 1999, Baron-Cohen & Wheelwright, 2004; Walker, 2005), dass Frauen und Mädchen eine besser ausgeprägte ToM aufweisen als Männer. Frauen und Mädchen sind sozial kompetenter und zeichnen sich durch prosoziales Verhalten und Empathiefähigkeit aus (Walker, 2005). Es zeigte sich in Untersuchungen, dass Mädchen bei *false belief*-Aufgaben, zur Überprüfung der ToM, deutlich besser abschneiden als Burschen (Cutting & Dunn, 1999, Charman et al., 2002, Wellmann et al., 2001).

Diese bisherigen Ergebnisse sind konsistent mit den Resultaten der vorliegenden Untersuchung, im Rahmen derer 1093 SchülerInnen der 12.Schulstufe getestet wurden. Es zeigte sich, dass Mädchen deutlich besser in den ToM-Aufgaben abschneiden als Burschen ($H_1(1)$). Relevant sind diesbezüglich auch die Ergebnisse des Wissenstests, der zur Überprüfung der Aktivierungswirksamkeit der Stereotype-Threat-Bedingungen vorgegeben wurde. Im Rahmen dieses Wissenstests wurde jeweils eine Fremd- und eine Selbsteinschätzungsaufgabe in Bezug auf die ToM vorgegeben. Die Ergebnisse weisen darauf hin, dass Frauen sowohl von den männlichen als auch von den weiblichen Schülern, unabhängig von den vier Stereotype Threat-Bedingungen, als sozial kompetenter und empathischer eingeschätzt wurden, was auf ein klares Stereotyp in Bezug auf die ToM zu Gunsten der Frauen hindeutet. Es ist interessant, dass keine der Stereotype Threat-Bedingungen diesbezüglich einen Einfluss auf die Fremdeinschätzung aufweist. In der Literatur gibt es zahlreiche Befunde dafür, dass Frauen in einer Testsituation in der Stereotype-Threat-aktivierenden Versuchsbedingung, in der auf geschlechtsspezifische Unterschiede zugunsten der Männer hingewiesen wird, schlechtere Leistungen als Männer und als ihre weiblichen Kollegen in einer neutralen Bedingung aufweisen, in der betont wird,

dass es keine geschlechtsspezifischen Fähigkeitsunterschiede gebe (Spencer et al., 1998; Spencer & Steele, 1999; Cadinu et al., 2002). In der Selbsteinschätzung zeigte sich, dass sich Mädchen in der ToM deutlich besser einschätzen als Burschen. Des Weiteren konnte ein signifikanter Unterschied in der Selbsteinschätzung der ToM in Abhängigkeit von den vier Stereotype Threat-Bedingungen festgestellt werden, der die erwartete Wirksamkeit der Aktivierungs- und Negativaktivierungsbedingung belegt. Diesbezüglich zeigte sich, wie auch Befunde aus der Literatur nahe legen (Steele, 1995; 1997; Spencer et al., 1999), dass die Aktivierung von Stereotype Threat einen mindernden Einfluss auf die Selbsteinschätzung und die Negativaktivierung, genau gegensätzlich, einen steigernden Einfluss auf die Selbstbeurteilung hat. Darüber hinaus kann eine signifikante Interaktion in der Selbsteinschätzung der ToM in Abhängigkeit vom Geschlecht und den Stereotype Threat-Bedingungen angenommen werden. Es konnte ein signifikanter geschlechtsspezifischer Unterschied zugunsten der Mädchen festgestellt werden. Mädchen schätzen sich, unabhängig von den Stereotype Threat-Bedingungen, durchwegs homogen ein. Bei der Selbsteinschätzung zeigte sich, wie bei der Fremdeinschätzung, dass die Stereotype Threat-Bedingungen keinen Einfluss auf die Beurteilungen der Mädchen haben. Interessanterweise ließen sich Burschen von der Aktivierungsbedingung beeinflussen, in der angeführt wurde, dass Frauen besser in Gedächtnisaufgaben und Männer besser in mathematischen und visuell-räumlichen Aufgaben seien. Insbesondere in dieser Bedingung schätzen sich Burschen wesentlich schlechter ein als Mädchen.

Die Prüfung der Haupthypothesen zur ToM ergab weitere interessante und auch überraschende Ergebnisse. In Bezug auf die vier Stereotype Threat-Bedingungen konnte zwar ein signifikanter Unterschied festgestellt werden ($H_1(3)$), jedoch fielen die Ergebnisse der paarweisen Vergleiche mittels Post-hoc-Tests nach Bonferroni nicht signifikant aus und somit war es nicht möglich zu identifizieren, zwischen welchen der Stereotype Threat-Bedingungen sich diese Unterschiede genau befinden. Es kann jedoch eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und den Stereotype Threat-Bedingungen angenommen werden ($H_1(4)$). Diesbezüglich konnte beobachtet werden, dass

Burschen in der Aktivierungs- Neutral- und Negativaktivierungsbedingung deutlich niedrigere Werte in den ToM-Aufgaben erzielen als Mädchen. Diese Niveauunterschiede verschwanden jedoch in der Aufklärungsbedingung, in der die Testpersonen über das Phänomen Stereotype Threat aufgeklärt werden. Mädchen erzielten grundsätzlich über alle Stereotype Threat-Bedingungen hinweg höhere Werte in den ToM-Aufgaben als Burschen. Es zeigte sich jedoch, dass Mädchen in der Neutralen Bedingung vergleichsweise die höchsten Werte erreichen. Dieses Ergebnis deckt sich mit Befunden aus der Literatur (Spencer et al., 1999), in denen bestätigt wurde, dass Frauen bessere Leistungen erzielen, sobald sie instruiert werden, dass es bezüglich bestimmter Fähigkeiten keine geschlechtsspezifischen Unterschiede gibt. Überraschenderweise weisen Mädchen in der Aktivierungs- und Negativaktivierungsbedingung annähernd gleich hohe Werte auf. Dieses Resultat entspricht nicht der ursprünglich angenommenen Erwartung, dass Frauen in der Negativaktivierung (Frauen sind besser in mathematischen und visuell-räumlichen Aufgaben) deutlich besser als in der Aktivierungsbedingung abschneiden würden. Diese Ergebnisse könnten darauf hindeuten, dass die Aktivierung von negativen Stereotypen und die Betonung geschlechtsspezifischer Unterschiede im Bereich der ToM und im Bewältigen der ToM-Aufgaben größere Auswirkungen und einen bedeutenderen Einfluss auf Burschen haben als auf Mädchen und diese diesbezüglich auch vulnerabler für die Bedrohung durch Stereotype sind als Mädchen. Es kann unter anderem als Hinweis auf das Bestehen eines Stereotyps zur ToM zugunsten der Frauen verstanden werden. Darüber hinaus könnte als eine Erklärung für die durchwegs homogene, gute Leistung der Mädchen in den ToM-Aufgaben über die vier Stereotype Threat-Bedingungen hinweg, der Ansatz der in der Literatur angeführten *multiplen sozialen Identitäten* dienen, zu dem es Untersuchungen von Schmader (2002) und Rydell und Beilock (2009) gibt. In diesen Studien wird thematisiert, dass gemäß der sozialen Identitätstheorie von Taifel und Turner (1986), jede Person von sich grundsätzlich ein positives Bild hat und danach bestrebt ist dieses Selbstbild zu erhalten. Die Annahme der multiplen sozialen Identitäten besagt, dass in einer Situation, in der zwei soziale Identitäten passend sind, jene verstärkt und zugänglich wird, die konsistent mit dem angestrebten,

positiven Selbstbild ist und einen selbstbewusstseinssteigenden Effekt hat, während der Zugriff auf die mit den negativen Stereotypen behafteten Identität gehemmt wird. In Bezug auf die vorliegenden Ergebnisse könnte dieser Ansatz dahingehend interpretiert werden, dass bei Mädchen im Rahmen der Bearbeitung der ToM-Aufgaben, annähernd unabhängig von den Stereotype Threat-Bedingungen, jene soziale Identität präsent und zugänglich war, die ein positives Selbstbild der Frau von sich selbst als fähige, kompetente, soziale und einfühlsame Person verkörperte und damit gleichzeitig den Zugriff zu der mit dem negativem Stereotyp der Schwäche und Unterlegenheit in mathematischen Fähigkeiten behafteten geschlechtsspezifischen Identität der Frau hemmt. Hierfür würden auch die Ergebnisse der Fremd- und Selbsteinschätzungen zur ToM im Rahmen des Wissenstests sprechen, wo sich Mädchen, wie bereits erwähnt, unabhängig von den Stereotype Threat-Bedingungen besser in der ToM einschätzen als Burschen.

In Bezug auf das Konstrukt des Risikoverhaltens konnten keine signifikanten Ergebnisse festgestellt werden. Es zeigte sich weder ein signifikanter Unterschied in der ToM in Abhängigkeit vom Risikoverhalten ($H_1(2)$) noch eine signifikante Wechselwirkung in der ToM in Abhängigkeit vom Geschlecht und dem Risikoverhalten ($H_1(5)$). Das bedeutet, dass im Zuge dieser Untersuchung keine Verbindungen zwischen dem Bewältigen der ToM-Aufgaben und den drei Ausprägungen des Risikoverhaltens hergestellt werden konnten. Das Risikoverhalten scheint somit keinen Einfluss auf das erfolgreiche Bearbeiten der ToM zu haben. Des Weiteren kann auch kein geschlechtsspezifischer Unterschied in der ToM im Zusammenhang mit den drei vorgegebenen Risikoinstruktionen beobachtet werden. Überraschenderweise konnte auch keine Wechselwirkung in der ToM in Abhängigkeit von den Stereotype Threat-Bedingungen und dem Risikoverhalten festzustellen ($H_1(6)$). In der Literatur lieferten Carr und Steele (2009) interessante Ergebnisse, die einen Zusammenhang zwischen Stereotype Threat, dem Verharren in ineffizienten Problemlösestrategien, dem *prevention focus* und risikoaversivem Verhalten aufdecken. Sie demonstrierten, dass die Bedrohung durch negative Stereotype zu einem Verharren in ineffizienten Problemlösestrategien führt und die Einnahme eines *prevention focus*, einer

vorsichtigen und bewahrenden Haltung begünstigt, die eingenommen wird, wenn sich eine Person Gedanken und Sorgen, um ihre Sicherheit macht. In weiterer Folge wird dadurch risikoaversives und vermeidendes Verhalten ausgelöst. Im Rahmen dieser Arbeit werden diese Zusammenhänge nicht bestätigt, obwohl die Annahme bestand, dass Unterschiede in den Risikobedingungen auch mit unterschiedlichen hohen ToM-Werten einhergehen würden. Des Weiteren wurde vermutet, dass Mädchen in der Aktivierungsbedingung und Ratebedingung schlechter in den ToM-Aufgaben abschneiden würden als in den anderen Bedingungen und als die Gruppe der Burschen. Darüber hinaus kann auch keine signifikante Interaktion höherer Ordnung in der ToM in Abhängigkeit vom Geschlecht, den Stereotype Threat-Bedingungen und dem Risikoverhalten angenommen werden ($H_1(7)$)

Ein weiteres Ziel dieser Arbeit war es, mögliche Zusammenhänge zwischen der ToM, dem Bewältigen der ToM-Aufgaben und dem Abschneiden in den sechs EMS-Untertests abzuleiten ($H_1(8-14)$). Diesbezüglich zeigen die Ergebnisse sehr kleine, positive signifikante Zusammenhänge zwischen der ToM und den EMS-ähnlichen Dimensionen (*Quantitative und formale Probleme, Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Diagramme und Tabellen, Figuren und Fakten lernen*). Es konnte beobachtet werden, dass höhere und bessere ToM-Leistungen auch mit höheren Werten in den einzelnen EMS-Untertests einhergehen. Eine Betrachtung der Zusammenhänge zwischen der ToM und dem Abschneiden in den einzelnen EMS-Untertests zeigte, dass im Speziellen vier Untertests (*Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Figuren und Fakten lernen*) identifiziert werden konnten, die den größten Erklärungswert an der ToM-Leistung ausmachten. Insbesondere der Untertest *Schlauchfiguren* scheint den relativ bedeutendsten Anteil für die Prognosefähigkeit im Rahmen der Bewältigung der ToM-Aufgaben aufzuweisen. Dieses Ergebnis ist dahingehend bedeutsam, da der Untertest *Schlauchfiguren* räumliches Raumvorstellungsvermögen erfasst, das zu den Stärken der Männer und Burschen zählt und einen der sieben EMS-Untertests darstellt, in denen Männer besser abschneiden als Frauen (Mallinger et al., 2009).

Des Weiteren stellte die geplante Studienwahl mit dem Fokus auf dem Wunsch nach einem später angestrebten Medizinstudium einen interessierenden und relevanten Aspekt im Rahmen dieser Studie dar. Die Untersuchung der ToM in Anhängigkeit vom Wunsch, Medizin zu studieren ergab jedoch kein signifikantes Ergebnis ($H_1(15)$). Dieses Ergebnis kann dahingehend interpretiert werden, dass die geplante Studienwahl keinen Einfluss auf das Bewältigen der ToM-Aufgaben hat. Interessanterweise kann eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und dem Wunsch, Medizin zu studieren festgestellt werden ($H_1(16)$). Die Betrachtung dieser Unterschiede erbrachte ein signifikantes Ergebnis für den Unterschied im ToM-Score innerhalb der Gruppe *ohne* Wunsch, Medizin zu studieren. Diesbezüglich ist festzuhalten, dass Mädchen innerhalb dieser Gruppe höhere Werte in der ToM erreichten als Burschen. Es zeigt sich tendenziell, dass Mädchen mit einer ausgeprägten ToM und höheren ToM-Werten nicht ihr geplantes Studium in dem Bereich der Medizin sehen. Das Ergebnis für die Gruppe der SchülerInnen mit dem Wunsch nach einem zukünftigen Medizinstudium war nicht signifikant, woraus sich ableiten lässt, dass kein geschlechtsspezifischer Unterschied im ToM-Score innerhalb dieser Gruppe angenommen werden kann.

Ein weiteres Anliegen dieser Untersuchung war es, das Risikoverhalten, zusätzlich zu den drei Risikobedingungen, mittels eines computerisierten Würfelspiels, dem *Game of Dice Task* (GDT), zu operationalisieren. Es sollten Unterschiede im Risikoverhalten, insbesondere in der Ausprägung der Risikobereitschaft und Risikofreude untersucht werden. Berichtete Ergebnisse aus der Literatur liefern Befunde, dass Männer risikoreicher agieren und eine höhere Ratetendenz aufweisen als Frauen (Ben-Shakar & Sinai, 1991; Byrnes, Miller & Schafer, 1999). Hinsichtlich des GDT zeigen Studienergebnisse, dass dieses objektive Verfahren dazu geeignet ist, Entscheidungen unter Risikobedingungen abzubilden (Brand et al. 2004; Brand et al. 2005, Eutenauer et al., 2009). Die Ergebnisse in der vorliegenden Arbeit können einen signifikanten Unterschied in der Risikobereitschaft, operationalisiert durch den GDT, in den instruierten Risikobedingungen identifizieren ($H_1(17)$). Die paarweisen Vergleiche mittels a-priori-Kontraste ergeben signifikante Unterschiede zwischen den Bedingungen

Raten und *ohne Instruktion*, sowie zwischen den Bedingungen *Raten* und *Sicher*. Hinsichtlich den Bedingungen *ohne Instruktion* und *Sicher* können keine signifikanten Ergebnisse berichtet werden. Bezüglich des beobachteten signifikanten Unterschieds zwischen der *Rateinstruktion* und der Bedingung *ohne Instruktion* ist zu festzuhalten, dass sich die Risikoinstruktion in diesem Fall als wirksam erwies und UntersuchungsteilnehmerInnen riskanter entschieden als in den anderen beiden Bedingungen. In der *Rateinstruktion* konnte im Vergleich zu den anderen Bedingungen der niedrigste Mittelwert festgestellt werden, was als Hinweis auf risikoreiches und risikofreudiges Entscheidungsverhalten gewertet werden kann, da im Rahmen des GDT niedrige Werte für risikoreiches Verhalten stehen. Es kann somit angenommen werden, dass TestteilnehmerInnen unter der Aufforderung zu *Raten* auch risikoreicher handelten. In der Literatur wird eine Entscheidung unter Risikobedingungen als eine Entscheidung unter Unsicherheit definiert, da es sich dabei um eine Entscheidung handelt, dessen Konsequenzen unsicher sind (Jungermann et al., 2005). In diesem Kontext kann man auch den Begriff des *Ratens* verstehen, da die Konsequenzen des *Ratens* ebenfalls unsicher sind und sowohl Gewinn als auch Verlust bringen können. In Bezug auf das Geschlecht konnte ein knapp nicht signifikanter Unterschied festgestellt werden ($H_1(18)$), was bedeutet, dass kein Unterschied in der Risikobereitschaft in Abhängigkeit vom Geschlecht und der Risikoinstruktion angenommen werden kann. Die Effektstärkenberechnung ergab einen geringen Effekt. Es kann lediglich eine Tendenz beobachtet werden. Dieses Ergebnis könnte darauf zurückzuführen sein, dass nicht alle UntersuchungsteilnehmerInnen den GDT durchgeführt haben und somit auch weniger SchülerInnen in die Berechnungen eingeschlossen werden konnten. Des Weiteren konnte auch keine Wechselwirkung in der Risikobereitschaft in Abhängigkeit vom Geschlecht und der Risikoinstruktion festgestellt werden ($H_1(19)$). Aus Zeitgründen konnte der GDT nicht allen UntersuchungsteilnehmerInnen vorgegeben werden, was zu den Einschränkungen und Kritikpunkten dieser Untersuchung überleitet. Der GDT wurde den SchülerInnen im Anschluss an die Testung vorgegeben und konnte aufgrund dessen nicht als Stör- bzw. Einflussvariable auf die Leistungen in den EMS-ähnliche Aufgaben und den ToM-Aufgaben behandelt werden, da dies nur bei

einer Vorgabe des GDT vor und nach der Testung möglich gewesen wäre. Infolgedessen ist die Aussagekraft und Verwertung der Ergebnisse des GDT etwas eingeschränkt. Des Weiteren ist anzuführen, dass die durchgeführte Testung nicht mit dem Originalsetting einer Aufnahmeprüfung gleichzusetzen ist, da die SchülerInnen keine Konsequenzen bezüglich des Bestehens oder Nicht-Bestehens der Testung zu erwarten hatten und somit auch nicht jenem Druck und Stress ausgesetzt waren wie potenzielle StudiumsanwärterInnen, diese Testung unter allen Umständen erfolgreich zu absolvieren.

Einen weiteren, zu diskutierenden Kritikpunkt betrifft die niedrigen Reliabilitäten der EMS-ähnlichen Untertests *Quantitative und formale Probleme, Medizinisch-naturwissenschaftliches Grundverständnis und Diagramme und Tabellen*. Die Berechnung der Reliabilitäten der Faktorstufenkombinationen erhöhte und verbesserte die Reliabilitäten geringfügig. Folgende Gründe stellen mögliche Erklärungen für die unbefriedigenden Ergebnisse dar:

1. Homogene Stichprobe

Grundsätzlich sind Reliabilitäten stichprobenabhängig. Bei homogenen Stichproben verhält sich ein Test weniger reliabel als bei heterogenen Stichproben (Bortz & Döring S.198, 2006) Im Rahmen dieser Untersuchung wurde eine relativ homogene Stichprobe von Jugendlichen im Alter von 17 bis 21 Jahren mit annähernd gleichem Bildungsniveau getestet.

2. EMS-ähnliche Aufgaben

Es war von den Testautoren intendiert, jedes Item so zu konstruieren, dass die *lokale stochastische Unabhängigkeit* gegeben ist. Unter *lokaler stochastischer Unabhängigkeit* versteht man die logische Unabhängigkeit zwischen den Items, was bedeutet, „dass die Beantwortung eines Items nicht eine vorhergehende Antwort auf ein anderes Item voraussetzen darf“ (Rost, S.96, 2004). Die Bewältigung der EMS-ähnlichen Aufgaben erfordert Wissen, was bedeutet, dass die TestteilnehmerInnen bei jeder Aufgabe vor eine neue Herausforderung gestellt werden und nachdenken müssen. Infolge dessen ist das Antwortverhalten der Testpersonen nicht konsistent. Die Reliabilitäten

geben Hinweise, dass die dahinterliegende latente Dimension nicht auf eine, sondern auf mehrere Fähigkeiten zurückzuführen ist (Rost, S.69, 2004).

Zeitbegrenzung

Für die Bearbeitung der einzelnen Untertests wurden Zeitbegrenzungen vorgegeben. Es handelt sich insofern um eine Testung mit einer Speedkomponente, aufgrund dessen inkonsistentes Lösungsverhalten mitverursacht wird. Im Zusammenhang mit der Speedkomponente könnten auch Positioneffekte eine Erklärung für die niedrigen Reliabilitäten sein. „Darunter versteht man die Veränderung der Schwierigkeit oder anderer Merkmale eines Items infolge seiner Platzierung im Test. Mit Positionseffekten ist besonders bei Items am Testanfang oder am Testende zu rechnen“ (Rost, S.383, 2004).

3. Itemanzahl pro Skala

Es wurden sechs EMS-ähnlichen Skalen vorgelegt, die ausserdem und jeweils aus weniger Items zusammengesetzt waren als die originalen EMS-Skalen. In Bezug auf die Reliabilität bedeutet dies, dass, bei einer geringeren Anzahl von Items einer Skala die Reliabilität niedriger ist (Rost, S.383, 2004)

4. Mehrdimensionalität der Skalen

Aus der vorliegenden Heterogenität der Items lässt sich ableiten, dass nicht nur eine Fähigkeit erfasst wird, sondern eine Mehrdimensionalität vermutet werden kann, was bedeutet, dass die Items mehrere Fähigkeiten konfundiert erfassen (Bortz & Döring, S.210, 2006).

5. Ernsthaftigkeit der Beantwortung

Des Weiteren ist anzumerken, dass sich die TestteilnehmerInnen nicht in einer realen Testsituation befanden, in der sie unter dem Druck des Bestehens oder Nicht-Bestehens standen. Es ist daher anzunehmen, dass die getesteten SchülerInnen nicht immer mit der notwendigen Ernsthaftigkeit an die Testung herangegangen sind und damit die Generalisierbarkeit und

Interpretationsmöglichkeiten der Ergebnisse begrenzt werden (Bortz & Döring, S.232, 2006).

Darüber hinaus ist anzumerken, dass eine verminderte Reliabilität auch mit einer niedrigeren Validität (Gültigkeit) einhergeht. Die Zuverlässigkeit kann als eine notwendige, aber nicht hinreichende Voraussetzung für die Validität gesehen werden (Rost, 2004)

Abschließend sollen noch Vorschläge und mögliche Richtungen für weitere Untersuchungen in diesem Feld erläutert werden. Die Thematik geschlechtsspezifischer Unterschiede im Zusammenhang mit der ToM und den Konstrukten Stereotype Threat und Risikoverhalten sollte im Rahmen einer Studie mit potenziellen AnwärterInnen für das Medizinstudium erfasst werden, um eine höhere Generalisierbarkeit der Ergebnisse zu erreichen und die realen Gegebenheiten im Originalsetting adäquater abzubilden. Es wäre bedeutsam, Unterschiede in der ToM zwischen MedizinstudentInnen und StudentInnen anderer Studienrichtungen zu untersuchen und Verbindungen zum Prüfungs- und Lernerfolg und -verhalten herzustellen. Des Weiteren stellen die Konstrukte Stereotype Threat und Risikoverhalten relevante Aspekte dar, die nachweislich Auswirkungen auf die Leistungen haben und dessen Wirkungsweise im Zusammenhang mit der ToM noch spezifischer untersucht werden sollte.

Kapitel 10. Zusammenfassung

Der Hauptgegenstand der vorliegenden Arbeit war die Untersuchung der ToM im Rahmen der Bewältigung EMS-ähnlicher Aufgaben in Abhängigkeit von dem Phänomen Stereotype Threat, dem Risikoverhalten und dem Geschlecht unter Berücksichtigung diesbezüglicher geschlechtsspezifischer Unterschiede. Diese empirische Arbeit stellt einen spezifischen Teil im Rahmen einer größer angelegten Studie dar, die von den StudentInnen Gabriele Hangl, Helene Lagger, Silvia Hameseder, Anita Teufl und Katrin Anzirk durchgeführt wurde. Im Zuge dieser Untersuchung behandelte jede der Kolleginnen einen anderen thematischen Schwerpunkt.

Das Anliegen dieser Diplomarbeit war es mögliche Zusammenhänge zwischen dem Bewältigen der ToM-Aufgaben, Stereotype Threat und dem Risikoverhalten abzuleiten und potenzielle wechselseitige Einflüsse dieser Konstrukte unter Berücksichtigung geschlechtsspezifischer Unterschiede zu untersuchen.

Bisherige Befunde aus der Literatur belegen, dass Frauen und Mädchen eine besser ausgeprägte ToM, die Kernkompetenzen wie zum Beispiel soziale Kognition, soziale Kompetenz, Perspektivenübernahme, Empathiefähigkeit und kognitive Flexibilität beinhaltet, aufweisen als Männer. Frauen und Mädchen sind sozial kompetenter und empathischer und werden auch als solches eingeschätzt (Bosacki et al., 1999, Baron-Cohen & Wheelwright, 2004; Walker, 2005). Des Weiteren zeigen Forschungsergebnisse, dass Mädchen in *false-belief*-Aufgaben, die zur Überprüfung der ToM-Fähigkeiten, basierend auf dem Konzept der Repräsentation falscher Überzeugungen eingesetzt werden, besser abschneiden als Burschen (Cutting & Dunn, 1999, Charman et al. 2002, Wellmann et al., 2001).

Es zeigte sich jedoch, dass die drei für diese empirische Arbeit relevanten und interessierenden Konstrukte der ToM, des Phänomens Stereotype Threat und des Risikoverhaltens unter Berücksichtigung geschlechtsspezifischer Unterschiede in der Literatur in diesem Zusammenhang bis dato noch nicht untersucht worden sind. Diesbezüglich gibt zwar Forschungsergebnisse, die belegen, dass Stereotype Threat ein Verharren in ineffizienten Problemlösestrategien begünstigt und dazu

führen kann, dass die betroffene Person eine vorsichtigeren und bewahrenden Haltung einnimmt, die zu risikoscheuem und risikoaversivem Verhalten führt, jedoch nicht unter Einbezug der ToM (Carr & Steele, 2009).

Im Rahmen der vorliegenden Untersuchung wurden insgesamt 1093 SchülerInnen (409 Burschen, 684 Mädchen) der 12. Schulstufe in Gymnasien in Wien, Niederösterreich und Burgenland getestet. Die UntersuchungsteilnehmerInnen waren zum Erhebungszeitpunkt im Mittel 17,57 Jahre alt. 17,5 % der getesteten SchülerInnen gaben an, ein späteres Medizinstudium anzustreben, während der Großteil mit 82,5% der UntersuchungsteilnehmerInnen diesbezüglich keinen Wunsch äußerte. Die Testung dauerte insgesamt 100 Minuten, in denen den SchülerInnen, die auf vier Stereotype Threat-Bedingungen und drei Risikobedingungen aufgeteilt wurden, sechs EMS-ähnliche Aufgabenblöcke mit begrenzter Bearbeitungszeit, ToM-Aufgaben, das Leistungsmotivationsinventar, die Skala zur Erfassung der Allgemeinen Selbstwirksamkeitserwartung und ein Computertest, der Game of Dice Task (GDT) vorgegeben wurden.

Die Ergebnisse zeigen, dass ein signifikanter Unterschied im Bewältigen der ToM-Aufgaben in Abhängigkeit vom Geschlecht festgestellt werden konnte ($H_1(1)$). Mädchen erreichten höhere Werte in den ToM-Aufgaben als ihre männlichen Schulkollegen. Des Weiteren konnte ein signifikanter Unterschied in der ToM in Abhängigkeit von den vier Stereotype Threat-Bedingungen angenommen werden ($H_1(3)$). Die genauere Betrachtung des Unterschieds mittels Post-hoc-Tests nach Bonferroni ergaben jedoch keine signifikanten Ergebnisse für die paarweisen Vergleiche. In Bezug auf das Risikoverhalten konnte kein signifikanter Unterschied im Bewältigen der ToM festgestellt werden ($H_1(2)$).

Es konnte jedoch eine signifikante Wechselwirkung in der ToM in Abhängigkeit vom Geschlecht und den vier Stereotype Threat-Bedingungen angenommen werden ($H_1(4)$). Diesbezüglich zeigte sich, dass Burschen in den Bedingungen *Aktivierung*, *Neutral* und *Negativ-Aktivierung* schlechter in den ToM-Aufgaben abschnitten als die Gruppe der Mädchen. Diese Niveauunterschiede verschwanden jedoch in der Aufklärungsbedingung, in der die TestteilnehmerInnen über das Phänomen aufgeklärt wurden. Mädchen erzielten

die relativ höchsten Werte in der Neutral-Bedingung, erreichten jedoch durchwegs homogen in allen Stereotype Threat-Bedingungen höhere ToM-Werte als ihre männlichen Schulkollegen. In Bezug auf das Risikoverhalten und das Geschlecht konnte keine signifikante Wechselwirkung nachgewiesen werden ($H_1(5)$). Des Weiteren konnte auch keine signifikante Wechselwirkung in der ToM in Abhängigkeit von den Stereotype Threat-Bedingungen und dem Risikoverhalten festgestellt werden ($H_1(6)$). Darüber hinaus fiel das Ergebnis für die Interaktion höherer Ordnung in der ToM in Abhängigkeit vom Geschlecht, den Stereotype Threat-Bedingungen und dem Risikoverhalten nicht signifikant aus ($H_1(7)$).

Hinsichtlich der Untersuchung des Zusammenhanges zwischen der ToM und dem Abschneiden in den sechs EMS-Untertests konnten sehr niedrige, positive signifikante Unterschiede festgestellt werden ($H_1(8-14)$). Es kann angenommen werden, dass höhere Werte in den ToM-Aufgaben auch mit höheren Werten in den einzelnen sechs EMS-Untertests einhergehen. Zusätzlich konnten vier Untertests (*Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Figuren und Fakten lernen*) mit Erklärungswert an der ToM identifiziert werden, wobei bei genauerer Betrachtung der Untertest *Schlauchfiguren* den bedeutendsten Anteil an der Prognosefähigkeit aufweist.

Die Prüfung der Nebenhypothesen zur ToM ergab für den Unterschied in der ToM in Abhängigkeit von dem Wunsch, Medizin zu studieren kein signifikantes Ergebnis ($H_1(15)$). Es konnte jedoch eine signifikante Interaktion in der ToM in Abhängigkeit vom Geschlecht und dem Wunsch, Medizin zu studieren festgestellt werden ($H_1(16)$). Diesbezüglich zeigten die paarweisen Vergleiche ein signifikantes Ergebnis für die Gruppe *ohne Wunsch, Medizin zu studieren*. Innerhalb dieser Gruppe erzielten die Mädchen höhere Werte in der ToM als die Burschen. In Bezug auf die Gruppe *Wunsch, Medizin zu studieren* konnte kein signifikantes Ergebnis nachgewiesen werden.

Des Weiteren konnte ein signifikanter Unterschied in der Risikobereitschaft, operationalisiert durch den GDT, in den instruierten Risikobedingungen festgestellt werden ($H_1(17)$). Die genauere Betrachtung der Unterschiede weist auf signifikante Unterschiede zwischen der *Rateinstruktion* und *ohne Instruktion* und

den Bedingungen *Raten* und *Sicher* hin. Hinsichtlich den Bedingungen *ohne Instruktion* und *Sicher* konnten keine signifikanten Ergebnisse festgestellt werden. Bezüglich des Geschlechts konnte eine Tendenz, aber kein signifikantes Ergebnis festgestellt werden ($H_1(18)$). Es kann somit kein Unterschied in der Risikobereitschaft in Abhängigkeit vom Geschlecht und der Risikoinstruktion angenommen werden. Darüber hinaus konnte keine signifikante Interaktion in der Risikobereitschaft in Abhängigkeit vom Geschlecht und der Risikoinstruktion beobachtet werden ($H_1(19)$). In der *Rateinstruktion* zeigte sich im Vergleich zu den anderen Bedingungen der niedrigste Mittelwert, was als Hinweis auf risikoreiches und risikofreudiges Entscheidungsverhalten gewertet werden kann, da im Rahmen des GDT niedrige Werte für risikoreiches Verhalten stehen.

11. Abstract

The aim of the study was the investigation of the *theory of mind* (ToM) in the course of attending EMS-tasks related to the phenomena *stereotype Threat*, *risk-behavior* and *sex*, considering gender-specific differences. *ToM* is the ability to attribute mental states to people as a way of understanding their feelings, thoughts and intentions in order to predict their actions (Förstl, 2007). *Stereotype threat*, the situation-specific threat of confirming negative stereotypes about one's group is activated in presence of negative stereotypes about this group and furthermore results in a reduced performance (Steele, 1997). In the course of the present study scholars (409 boys and 684 girls) aged between 16 and 21 were tested. Four different stereotype threat (activation, neutral, clarification: description of the phenomena and negative-activation) and three risk-behavior conditions, (guessing, no instruction, being sure) were maintained to examine the influence of stereotype threat and risk-behavior while attending six different EMS- tasks and *ToM*- tasks. The results showed a significant difference between boys and girls, favoring girls concerning the performance on *ToM* tasks. Furthermore a significant interaction between stereotype threat conditions and sex could be identified. Boys performed worse than girls in *ToM*- tasks in three out of four conditions. These differences seem to disappear in the condition clarification. Girls performed in every condition nearly equally good. Regarding the risk-behavior conditions, there could be no significant difference related to *ToM* or stereotype threat assessed. Concerning possible relations between *ToM* and the performance in EMS-tasks, results indicate small and positive relations between *ToM* and the six EMS-tasks. It could be assumed that higher *ToM*-scores go along with higher scores in every single EMS-task. Another interest of this study was to examine if there is a relation between *ToM*, the request of earning a medical degree considering sex-specific differences. Referring to this, results showed significant differences between boys and girls in *ToM*-tasks related to the group with *no request of earning a medical degree*. Within this group girls performed significantly better in *ToM*-tasks than boys.

Literaturverzeichnis:

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C.M. & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in Stereotype Threat. *Journal of Experimental Social Psychology* 35, 29-46.
- Aronson, J., Quinn, D. M. & Spencer, S.J. (1998). Stereotype threat and academic underperformance of minorities and women. In J.K. Swim & C. Stangor (Eds.). *Prejudice: The target's perspective* (pp.83-103). San Diego: Academic Press.
- Aronson, J., Fried, C.B. & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, 38, 113–125.
- Aronson, E., Wilson, T. D. & Akert, R. M. (2008). *Sozialpsychologie (sechste, aktualisierte Auflage)*. München: Pearson Studium.
- Astington, J. W. & Jenkins, J. (1999). A longitudinal study of the relation between language and theory of mind development. *Developmental Psychology*, 35, 1311-1320.
- Astington, J.W. (2000). *Wie Kinder das Denken entdecken*. München: Ernst Reinhardt Verlag.
- Astington, J. W. & Dack, L. A. (2008). Theory of Mind. In: Haith, M.M. & Benson, J. B. (Eds.). *Encyclopedia of Infant and Early Childhood Development*. 3, Elsevier Inc. British Library Cataloguing.
- Baron-Cohen, S. & Hammer, J. (1997). Is autism an extreme form of male brain? *Advances in Infancy Research*, 11, 193-217.

- Baron-Cohen, S., Bolton, P., Wheelwright, S., Scahill, V., Short, L., Mead, G. & Smith, A. (1998). Autism occurs more often in families of physicists, engineers, and mathematicians. *Autism*, 2, 296-301.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6, 248-254.
- Baron-Cohen, S. & Wheelwright, S. (2004). The Empathy Quotient: An investigation of adults with Asperger Syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34, 163-175.
- Baron-Cohen, C., Knickmeyer, R. C. & Belmonte, M.K. (2005). Sex differences in the brain: Implications for explaining autism. *Science*, 310, 819-823.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2011). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin, Heidelberg: Springer.
- Becker, H.S. (1981): *Außenseiter. Zur Soziologie abweichenden Verhaltens*. Frankfurt am Main: Fischer Verlag
- Ben-Shakhar, G. & Sinai, Y. (1991). Gender differences in multiple choice tests: the role of differential guessing tendencies. *Journal of Educational Measurement*, 28, 23-35.
- Ben-Zeev, T., Fein, S. & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, 41, 174-181.
- Bischof-Köhler, D. (2000). *Kinder auf Zeitreise: Theory of Mind, Zeitverständnis und Handlungsorganisation*. Bern: Hans Huber.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation*. Berlin, Heidelberg: Springer.

Bosacki, S., University, D. & Astington, J. W. (1999). Theory of mind in preadolescence: Relations between social understanding and social competence. *Social Development, 8*, 237-255.

Bosco, A., Longoni, A. M. & Vecchi, T. (2004). Gender effects in spatial orientation: Cognitive profiles and mental strategies. *Applied Cognitive Psychology, 18*, 519-532.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*, 247-255.

Brand, M., Fujiwara, E., Borsutzky, S., Markowitsch, H.J., Kalbe, E. & Kessler, J. (2005). Decision-making deficits of Korsakoff patients in a new gambling task with explicit rules: Associations with executive functions. *Neuropsychology, 19*, 267-277.

Brand, M., Labudda, K., Kalbe, E., Hilker, R., Emmans, D., Fuchs, G. et al. (2004). Decision-making impairments in patients with Parkinson's disease. *Behavioural Neurology, 15*, 77-85.

Breakwell, G. M. (2007). *The Psychology of Risk*. Cambridge: University Press.

Brown, R. P. & Pinel, E. C. (2003). Stigma on mind: individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39*, 626-633.

Byrnes, J. P., Miller, D. C. & Schafer, W.D. (1999). Gender differences in risk taking: a meta-analysis. *Psychological Bulletin, 125*, 367-383.

Cadinu, M., Maass, A., Rosabianca, A. & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science, 16*, 572-578.

- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L. & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, 33, 267-285.
- Carr, P. B. & Steele C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, 45, 853-859.
- Cattell, R. B. & Cattell, A. K. S. (1963). *Test of "g": Culture Fair Scale 3, Form A*. Champaign, IL: Institute for Personality Testing.
- Charman, T., Ruffmann, T. & Clements, W. (2002). Is there a gender difference in false belief development. *Social Development*, 11, 1-10.
- Cutting, A. & Dunn, J. (1999). Theory of Mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, 70, 853-865.
- Davis, M.H. & Franzoi, S.L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality*, 25, 70-87.
- Deaux, K. & Emsweiler, T. (1974). Explanations of successful performance of sex-linked tasks: What is skill for male is luck for the female. *Journal of Personality and Social Psychology*, 29, 80-85.
- Eisenführ, F. & Weber, M. (2003). *Rationales Entscheiden*. Heidelberg: Springer.
- Eutenauer, F., Schaefer, F., Stuermer, R., Boucsein, W., Timmermann, L., Barbe, M. T., Ebersbach, G., Otto, J., Kessler, J. & Kalbe E. (2009). Dissociation of decision-making under ambiguity and decision-making under risk in patients with Parkinson's disease: A neuropsychological and psychophysiological study. *Neuropsychologia*, 47, 2882-2890.
- Fenigstein, A., Scheier, M.F. & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43, 522-527.

- Ferstl, E. C. (2007). Theory of Mind und Kommunikation: Zwei Seiten derselben Medaille?. In: Förstl, H. (Hrsg.). *Theory of mind: Neurobiologie und Psychologie sozialen Verhaltens*. Heidelberg: Springer.
- Field, A. (2005). *Discovering Statistics Using SPSS (second Edition)*. London: Sage Publications.
- Fiske, S.T. & Depret, E. (1996). Control, interdependence, and power: Understanding social cognition in its social context. *European Review of Social Psychology*, 7, 31-61.
- Flavell, J.H., Green, F.L, Flavell, E.R & Grossman, J.B. 1997). The development of children`s knowledge about inner speech. *Child Development*, 68, 39-47.
- Fonagy, P., Steele, H., Moran, G., Steele, M. & Higgitt, A. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal*, 13, 200-217.
- Förstl, H. (2007). *Theory of mind: Neurobiologie und Psychologie sozialen Verhaltens*. Heidelberg: Springer.
- Frischenschlager, O., Mitterauer, L. & Haidinger, G. (2005). Leistungsfaktoren als potenzielle Auswahlkriterien im Medizinstudium. *Zeitschrift für Hochschuldidakik*, 6, 34-41.
- Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12(2), 229-246.
- Halpern, D. F. & Tan, U. (2001). Stereotypes and Steroids: Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cognition*, 45, 392-414.
- Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science*, 13, 135–139.

Hänsgen, K.-D. & Spicher, B. (2007): *EMS Eignungstest für das Medizinstudium 2007- Berichte des Zentrums für Testentwicklung, 13*. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.

Happé, F. G. E. (1995). The role of age and verbal ability in the theory –of- mind task performance of subjects with autism. *Child Development, 66*, 843-855)

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280–1300.

Hogrefe, G. J., Wimmer, H. & Perner J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 53*, 567-582.

Huang, J. (1993). An investigation of gender differences in cognitive abilities among Chinese high school students. *Personality & Individual Differences, 75*, 717-719.

Huber, O. (2004). *Entscheiden unter Risiko: Aktive Risiko-Entschärfung*.

ITB Consulting (2008). *Test für medizinische Studiengänge I: Originalversion I des TMS (5.Auflage)*. Göttingen: Hogrefe

ITB Consulting (2008). *Test für medizinische Studiengänge II: Originalversion II des TMS (5. Auflage)*. Göttingen: Hogrefe.

Jenkins, J. M. & Astington, J.W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology, 32*, 70-78.

Jungermann, H., Pfister, H.-R. & Fischer, K. (2005). *Die Psychologie der Entscheidung: Eine Einführung*. München: Spektrum Akademischer Verlag

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.

- Kimura, D. & Hampson, E. (1994). Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Current Directions in Psychological Science*, 3, 57-61.
- Kit, K. A., Tuokko, H. A. & Mateer, C.A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychological Review*, 18, 132-148.
- Kirchler, E. M. (2003). *Wirtschaftspsychologie. Grundlagen und Anwendungsfelder der ökonomischen Psychologie*. Göttingen: Hogrefe.
- Kühberger, A. (1994). Risiko und Unsicherheit: Zum Nutzen des Subjective Expected Utility-Models. *Psychologische Rundschau*, 45, 3-23.
- Laux, H. (2003). *Entscheidungstheorie*. Heidelberg: Springer
- Lawton, C. A., & Hatcher, D. W. (2005). Gender differences in integration of images in visuospatial memory. *Sex Roles*, 53(9-10), 717-725.
- Levy, B. & Langer, E. (1994). Aging free from negative stereotypes: Successful memory in China among the American deaf. *Journal of Personality and Social Psychology*, 66, 989-997.
- Lippmann, W. (1992). *Public Opinion*. New York: Free Press.
- Lowe, P. A., Mayfield, J. W., & Reynolds, C. R. (2003). Gender differences in memory test performance among children and adolescents. *Archives of Clinical Neuropsychology*, 18, 865-878.
- Maccoby, E. E. & Jacklin, C. N. (1974). *The Psychology of sex differences*. Stanford: Stanford University Press.
- Mallinger, R., Holzbaur, C., Dietrich, M., Heidegger, M., Zeller, J., Hänsgen, K.-D. & Spicher, B. (2009). *EMS Eignungstest für das Medizinstudium in Österreich. Bericht zum EMS Innsbruck und Wien*. Freiburg, Schweiz: Universität, Zentrum für Testentwicklung und Diagnostik, Departement für Psychologie.

Marx, D. M. & Stapel, D. A. (2006). It's all in the timing: Measuring emotional reactions to stereotype threat before and after taking a test. *European Journal of Social Psychology*, 36, 687-698.

O'Connor, T. G. & Hirsch, N. (1999). Intra-individual differences and relationship-specificity of mentalising in early adolescence. *Social Development*, 8, 256-274.

Oyserman, D., Uskul, A. K., Yoder, N., Nesse, R. M. & Williams, D. R. (2007). Unfair treatment and self-regulatory focus. *Journal of Experimental Social Psychology*, 43, 505-512.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, 76, 114-128.

Ploog, D. (2007). Ich, der andere und mein Wille. Anmerkungen zur Theory of Mind. In: Förstl, H. (Hrsg.). *Theory of mind: Neurobiologie und Psychologie sozialen Verhaltens*. Heidelberg: Springer.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *The Behavioral and Brain Sciences*, 4, 515-526.

Rost, J. (2004). *Testtheorie-Testkonstruktion*. Bern: Hans Huber.

Rubio, V. J., Hernández, J. M., Zaldivar, F., Márquez, O. & Santacreu, J. (2010). Can we predict risk-taking behavior? Two behavioral tests for predicting guessing tendencies in a multiple choice test. *European Journal of Psychological Assessment*, 26, 87-94.

Rydell, R. J., Beilock, S. L. & McConnell, A. R. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology*, 96, 949-966.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38, 194-201.

- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat on performance. *Psychological Review*, *115*, 336-356.
- Schuler, H. & Prochaska, M. (2001). *Leistungsmotivationsinventar (LMI)*. Göttingen: Hogrefe.
- Schwarzer, R. & Jerusalem, M. (Hrsg.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.
- Sodian, B. (2002). Entwicklung begrifflichen Wissens. In: Oerter, R. & Montada, L. (Hrsg.). *Entwicklungspsychologie* (fünfte, vollständig überarbeitete Auflage). Weinheim: Beltz Verlag.
- Sodian, B. (2007). Entwicklung der Theory of Mind in der Kindheit. In: Förstl, H. (Hrsg.). *Theory of mind: Neurobiologie und Psychologie sozialen Verhaltens*. Heidelberg: Springer.
- Spence, J. T., Helmreich, R. L. & Sawin, L. L. (1980). The male-female relations questionnaire: A self-report inventory of sex role behaviors and preferences and its relationships to masculine and feminine personality traits, sex role attitudes, and other measures. *JSAS Catalog of Selected Documents in Psychology*, *10*, 87-88.
- Spencer, S.J., Steele, C.M. & Quinn, D.M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4-28.
- Spiel, C., Schober, B. & Litzenberger, M. (2007). *Evaluation der Eignungstests für das Medizinstudium in Österreich*. Wien: Universität, Fakultät für Psychologie.
- Steele, C.M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance. *American Psychologist*, *52*, 613-629.

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811.

Tajfel, H., Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Augustin (Eds.), *The psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.

Vandenberg, S. G. & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual & Motor Skills*, 47, 599-601.

Walker, S. (2005). Gender differences in the relationship between young children`s peer-related social competence and individual differences in theory of mind. *The Journal of Genetic Psychology*, 166, 297-312.

Willinger, U., Schmöger, M., Müller, C. & Auff, E. *Theory of Mind-Stories*. In Vorbereitung.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children`s understanding of deception. *Cognition*, 13, 103-128.

Online-Quellen:

.(<http://www.eignungstest-medizin.at/cms/>(25.04.2011).

(<http://www.unifr.ch/ztd/ems/EMSaufbau.pdf>, 25.04.11).

(<http://ug.manz.at>, 30.08.2011)

Tabellenverzeichnis:

Tabelle 1: 10 EMS-Untertests und die dadurch erfassten Fähigkeiten.....	6
Tabelle 2: Untersuchungsdesign	70
Tabelle 3: Reliabilitäten nach Cronbachs Alpha (α) für die EMS-Untertests von 1998-2007 nach Hänsgen und Spicher (2007).....	77
Tabelle 4: Reliabilitäten der sechs eingesetzten EMS-Untertests (Hänsgen & Spicher, 2007).....	77
Tabelle 5: Häufigkeiten in Abhängigkeit von Geschlecht.....	85
Tabelle 6: Signifikanzbeurteilung der Geschlechterverteilung	86
Tabelle 7: Verteilung der SchülerInnen auf die vier Stereotype Threat-Bedingungen	87
Tabelle 8: Signifikanzprüfung der Verteilung auf die Stereotype Threat Bedingungen	87
Tabelle 9: Kontingenztafel Geschlecht und Stereotype Threat	88
Tabelle 10: Verteilung der SchülerInnen auf die drei Risikobedingungen	89
Tabelle 11: Signifikanzprüfung der Geschlechterverteilung in den Risikobedingungen	89
Tabelle 12: Kontingenztafel Geschlecht und Risikoverhalten	90
Tabelle 13: Kontingenztafel Stereotype Threat und Risikoverhalten.....	90
Tabelle 14: Häufigkeiten der Altersstufen.....	92
Tabelle 15: Signifikanzprüfungsprüfung Alter.....	93
Tabelle 16: Häufigkeiten und Anteilswerte der besuchten Schulzweige	94
Tabelle 17: Signifikanzprüfung Schulzweige.....	94
Tabelle 18: Häufigkeiten Geschlecht nach Wunsch Medizinstudium.....	95
Tabelle 19: Signifikanzprüfung Wunsch nach Medizinstudium	96
Tabelle 20: Kontingenztafel Geschlecht und Wunsch Medizinstudium	96
Tabelle 21: Kontingenztafel nach Wunsch Medizinstudium und EMS-Vorbereitung	97
Tabelle 22: Reliabilitäten nach Cronbachs Alpha der sechs EMS-Untertests....	100
Tabelle 23: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Untertest Quantitative und formale Probleme.....	101

Tabelle 24 Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Schlauchfiguren	101
Tabelle 25 Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Untertest Medizinisch-naturwissenschaftliches Grundverständnis	102
Tabelle 26: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Diagramme und Tabellen	102
Tabelle 27: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemsanzahl) EMS- ähnlicher Untertest Figuren Lernen	103
Tabelle 28: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Fakten lernen	103
Tabelle 29: Häufigkeiten Figuren lernen N= 1084.....	105
Tabelle 30: Kontingenztafel Stereotype Threat-Bedingungen und WT Figuren lernen (N= 1084).....	105
Tabelle 31: Häufigkeiten WT Fakten lernen N= 1082	107
Tabelle 32: Kontingenztafel Stereotype Threat-Bedingungen und WT Fakten lernen (N = 1082).....	107
Tabelle 33: Häufigkeiten WT Med.-naturwissenschaftl. GV N= 1080.....	108
Tabelle 34: Kontingenztafel Stereotype Threat-Bedingungen und WT Med.-nat.-wissenschaftl. GV	109
Tabelle 35;Häufigkeiten WT Diagramme und Tabellen N= 1080	110
Tabelle 36: Kontingenztafel Stereotype Threat-Bedingungen und WT Diagramme und Tabellen	110
Tabelle 37: Häufigkeiten WT Quantitative und formale Probleme N= 1080.....	112
Tabelle 38: Kontingenztafel Stereotype Threat-Bedingungen und WT Quantitative und formale Probleme (N = 1080)	112
Tabelle 39: Häufigkeiten WT Schlauchfiguren N=1080	113
Tabelle 40: Kontingenztafel Stereotype Threat-Bedingungen und WT Quantitative und formale Probleme (N = 1080)	114
Tabelle 41: Signifikanzbeurteilungen der Fremdeinschätzungsaufgaben	115
Tabelle 42 Statistische Kennwerte der Aktivierungs-Bedingung (A)	116
Tabelle 43:: Statistische Kennwerte der Neutral-Bedingung (B)	116

Tabelle 44: Statistische Kennwerte der Aufklärungs-Bedingung (C)	117
Tabelle 45: Statistische Kennwerte der Negativ-Aktivierungs-Bedingung (D) ..	117
Tabelle 46: Statistische Kennwerte der Selbsteinschätzungsaufgaben.....	118
Tabelle 47: : Selbsteinschätzungen WT.....	121
Tabelle 48: Statistische Kennwerte ToM-Selbsteinschätzungsaufgabe	122
Tabelle 49: ANOVA ToM Selbsteinschätzungsaufgabe	123
Tabelle 50: Post hoc-Tests nach Bonferroni.....	124
Tabelle 51: statistische Kennwerte der EMS-Untertests gesamt	127
Tabelle 52: Statistik Quantitative und formale Probleme.....	128
Tabelle 53: Häufigkeiten UT Quantitative und formale Probleme.....	128
Tabelle 54: statistische Kennwerte UT Schlauchfiguren.....	129
Tabelle 55: Häufigkeiten und Prozentwerte UT Schlauchfiguren.....	130
Tabelle 56: Häufigkeiten UT Medizinisch-naturwissenschaftliches Grundverständnis	132
Tabelle 57: Häufigkeiten UT Diagramme und Tabellen	133
Tabelle 58: statistische Kennwerte UT Fakten und Figuren.....	134
Tabelle 59: Häufigkeiten UT Figuren lernen.....	135
Tabelle 60: Häufigkeiten UT Fakten lernen	136
Tabelle 61: Statistische Kennwerte zum GDT	137
Tabelle 62: gültige Fälle	138
Tabelle 63: Levene-Test	139
Tabelle 64: Statistische Kennwerte Geschlecht und Stereotype Threat Bedingungen	140
Tabelle 65: Ergebnisse der ANOVA	142
Tabelle 66: Interkorrelationsmatrix der Koeffizienten der Produkt- Momentkorrelation ToM mit EMS-ähnlichen Aufgaben(n = 1085)	146
Tabelle 67: in das Modell aufgenommene Variablen	147
Tabelle 68: aus dem Modell ausgeschlossene Variablen	148
Tabelle 69: Statistische Kennwerte.....	148
Tabelle 70 :Welch-Test.....	149
Tabelle 71 Statistische Kennwerte Geschlecht und Wunsch nach Medizinstudium	150

Tabelle 72. ANOVA Geschlecht und Wunsch nach Medizinstudium	150
Tabelle 73: Gruppenstatistiken kein Wunsch nach Medizinstudium	152
Tabelle 74: t-Test: Gruppe ohne Wunsch nach Medizinstudium	152
Tabelle 75: Gruppenstatistiken Wunsch nach Medizinstudium	153
Tabelle 76: t-Test: Gruppe Wunsch nach einem Medizinstudium	153
Tabelle 77: Statistik Risikobedingungen in N = 471	154
Tabelle 78: ANOVA ns.GDT(Risikofreude), Geschlecht und Risikoverhalten..	155

Abbildungsverzeichnis:

Abbildung 1: Beispielitem 1- Fremdeinschätzung	75
Abbildung 2: Beispielitem 2- Fremdeinschätzung ToM	75
Abbildung 3: Beispielitem 3 – Selbsteinschätzung	76
Abbildung 4: Selbsteinschätzung ToM.....	76
Abbildung 5: Im Büro-zweikategorielles Antwortformat	82
Abbildung 6: Briefverkehr – dreikategorielles Antwortformat	83
Abbildung 7: In der Werbeagentur- vierkategorielles Antwortformat	83
Abbildung 8: Prozentwerte der Geschlechter in Prozent.....	86
Abbildung 9: Häufigkeitsverteilung (Anteilswerte in Prozent) der Altersstufen (1086 gültige Fälle).....	92
Abbildung 10: Wunsch Medizinstudium nach dem Geschlecht.....	95
Abbildung 11: Interaktionsdiagramm ToM-Aufgabe.....	125
Abbildung 12: Interaktionsdiagramm Diagramme und Tabellen	126
Abbildung 13: Verteilung Quantitative und formale Probleme.....	129
Abbildung 14: Verteilung Schlauchfiguren in der Gesamtstichprobe.....	130
Abbildung 15: Verteilung UT Medizinisch-naturwissenschaftliches GV	132
Abbildung 16: Verteilung UT Diagramme und Tabellen.....	134
Abbildung 17: Verteilung UT Figuren lernen	135
Abbildung 18: Verteilung UT Fakten lernen	136
Abbildung 19: Verteilung GDT-Nettoscore	137
Abbildung 20: Histogramm zum ToM-Score.....	138
Abbildung 21: Interaktionsdiagramm der WW Stereotype Threat und Geschlecht	144
Abbildung 22: Interaktionsdiagramm für ToM in Abhängigkeit vom Geschlecht und Wunsch, nach Medizinstudium.....	151
Abbildung 23: NS GDT- Score Risikofreude (NS GDT-Score) in Abhängigkeit vom Risikoverhalten und Geschlecht	156

Anhang:

A1. Elternbrief

Liebe Eltern, liebe Mütter und Väter!

In der Schule, die ihr Kind derzeit besucht, soll eine wissenschaftliche Untersuchung durchgeführt werden, in der es darum geht, die sozialen und biologischen Einflußfaktoren auf die Leistungsfähigkeit während Testsituationen zu untersuchen und Unterschiede zwischen männlichen und weiblichen Jugendlichen zu beurteilen. Die Ergebnisse sollen im Weiteren helfen, Eignungstests für beide Geschlechter fair zu gestalten.

Die Untersuchung wird im Rahmen von zwei Unterrichtsstunden an der Schule durchgeführt.

Selbstverständlich stehen dabei die Bedürfnisse ihres Kind im Vordergrund!

Diese Studie wird im Rahmen unserer Diplomarbeit unter der Betreuung von Frau Ao. Univ. Prof. Mag. Dr. Ulrike Willinger am psychologischen Institut der Universität Wien durchgeführt. Alle Informationen und Ergebnisse werden selbstverständlich vertraulich behandelt.

Bei Einverständnis der Teilnahme Ihres Kindes an der Untersuchung bitte ich Sie, diese Mitteilung zu unterschreiben und Ihrem Kind in die Schule mitzugeben.

Herzlichen Dank für ihre Unterstützung!

Mit freundlichen Grüßen

Katrin Anzirk, Silvia Hameseder, Gabriele Hangl, Helena Lagger und Anita Teufl

Ich erkläre mich damit einverstanden, dass mein Kind,

....., geboren am (Geburtsdatum des Kindes), an der Untersuchung teilnimmt.

Klasse des Kindes:

Datum und Unterschrift des
Erziehungsberechtigten:

A2. Information für LehrerInnen

Liebe Direktion, liebe Lehrerinnen und Lehrer!

Informationen zur Studie zum Medizineinstiegstest

Der Eignungstest für das Medizinstudium (EMS; Hänsgen, 1998) wird seit 2006 von der Medizinischen Universität Wien und der Medizinischen Universität Innsbruck eingesetzt. Er erfreut sich in der Schweiz und Deutschland großer Beliebtheit. Der EMS besteht aus 10 Subtests, die unterschiedliche kognitive Leistungen von den KandidatInnen verlangen. Die Analyse der geschlechtsspezifischen Ergebnisse im EMS zusammenfassend für 2006 bis 2008 ergibt für Österreich folgendes Bild: In den Subtests „Quantitative und formale Probleme“, „Schlauchfiguren“, „Textverständnis“, „Planen und Organisieren“, „Medizinisch-naturwissenschaftliches Grundverständnis“, „Diagramme und Tabellen“, „Konzentriertes und sorgfältiges Arbeiten“ und „Muster Zuordnen“ werden signifikante Leistungsunterschiede zugunsten der männlichen Kandidaten ausgewiesen. Frauen schneiden signifikant besser in den Subtests „Fakten lernen“ und „Figuren lernen“ ab.

Ähnliche geschlechtsspezifische Unterschiede von KandidatInnen, die an einer der beiden österreichischen medizinischen Universitäten den EMS absolviert haben, lassen sich in der Schweiz ebenso replizieren.

Eine Vielzahl von Studien belegt relative Leistungsunterschiede zwischen den Geschlechtern in kognitiven Aufgaben (u.a. Kimura, 2002). Als Ursachen dafür werden sowohl soziale als auch biologische Faktoren diskutiert. So konnte etwa gezeigt werden, dass die Wahrnehmung und Interpretation der Testsituation an sich einen wesentlichen Einfluss auf das Testergebnis der jeweiligen Testperson haben kann (Hausmann, 2007).

In diesem Zusammenhang wurde der Begriff „stereotype threat“ (Steele, 1997) geprägt. Dieses Konstrukt beschreibt ein Phänomen, das die eigene kognitive Leistungsfähigkeit negativ beeinflussen kann. Werden einer Personengruppe negative Stereotype gesellschaftlich zugesprochen (z.B. „Frauen tun sich schwerer, räumliche Aufgaben zu lösen“) und wird die Identifikation mit der stereotypisierten Gruppe aktiviert, kann ein „stereotype threat“ ausgelöst werden und Auswirkungen auf die eigene Leistungsfähigkeit haben (Hausmann, 2007).

Interessanterweise kann eine Aufklärung über dieses Phänomen bei Versuchspersonen dazu führen, dass sich der Effekt des „stereotype threat“ verringert bzw. nicht mehr zeigt.

Die Ergebnisse unterschiedlicher Studien zeigen auf, dass die Frauen der Kontrollgruppe, die keinem „stereotype threat“ ausgesetzt wurden, durchwegs bessere Resultate erbrachten als die Frauen der Versuchsgruppe, die mit „stereotype threat“ konfrontiert wurden (Kit, Tuokko & Mateer, 2008). Weiters konnte gezeigt werden, dass Frauen, die in einer stereotype threat reduzierten Bedingung getestet wurden, gleichgute Ergebnisse bezüglich ihrer mathematischen Testleistung erzielten wie ihre männlichen Kollegen (u.a. Keller and Dauheimer 2003).

Eine geringere Anzahl von Studien beschäftigte sich mit der Tendenz der Risikobereitschaft einer Testsituation (u.a. Ben-Shakhar & Sinai, 1991). Ben-Shakhar & Sinai (1991) untersuchten insbesondere Geschlechtsunterschiede bei der Beantwortung von Multiple-Choice-Tests unter der besonderen Berücksichtigung der Tendenz, zu raten und Items auszulassen. Es konnte festgestellt werden, dass Männer weniger dazu tendieren Items auszulassen als Frauen. Die Unterschiede zeigten sich stärker in kognitiven Aufgaben, die männliche Vorteile aufwiesen. Jedoch können die männlichen Vorteile nicht vollständig die Unterschiede beim Auslassen der Items erklären, da die Männer auch in kognitiven Aufgaben mit einer weiblichen Überlegenheit weniger Items ausließen. Weiters konnte gezeigt werden, dass die Testpersonen zum Auslassen mancher Items tendieren, auch wenn eine Instruktion zum Raten besteht, und diese Auslassungen geschlechtsbezogen sind (Ben-Shakhar & Sinai, 1991).

Ein weiterer interessanter Aspekt, der innerhalb der Studie erhoben werden soll ist, ob ein Zusammenhang zwischen einerseits dem Phänomen des Stereotype Threat und der Theory of Mind und andererseits der Risikobereitschaft bzw des Risikoverhaltens und der Theory of Mind besteht. Unter dem Begriff der **Theory of Mind** (ToM) wird sowohl in der Psychologie als auch in den anderen Kognitionswissenschaften die Fähigkeit verstanden, eine Annahme über Bewusstseinsvorgänge in anderen Personen vorzunehmen und diese in der eigenen Person zu erkennen, also Gefühle, Bedürfnisse, Ideen, Absichten, Erwartungen und Meinungen zu vermuten. In der Literatur gibt es bis zu dem jetzigen Zeitpunkt keine Studien und wissenschaftlichen Erkenntnisse, die diese Fragestellung untersucht oder berücksichtigt hätten.

Das Ziel der vorliegenden Studie ist ein multivariater Ansatz zur Berücksichtigung von sozialen Faktoren (Stereotype Threat, Risikobereitschaft) zur Klärung von geschlechtsspezifischen Unterschieden in kognitiven Leistungen (= EMS-ähnliche

Aufgaben und eine sozial-kognitive Aufgabe) unter Berücksichtigung von „mediating variables“ (Selbstwirksamkeit, Leistungsmotivation).

Hauptthese:

Es gibt einen signifikanten Unterschied zwischen Männern und Frauen, in den 4 „Stereotype Threat-Bedingungen“ und in den 3 „Risk-Behavior-Bedingungen“ im Bewältigen von kognitiven Aufgaben (EMS-ähnlicher Test, Theory of Mind) unter der Berücksichtigung des Hormonstatus, der individuellen Ängstlichkeit, Selbstwirksamkeit und Leistungsmotivation.

Untersuchungsablauf:

Zur Messung des Stereotype Threat werden vier verschiedene Untersuchungsbedingungen durchgeführt. In der ersten Bedingung werden die SchülerInnen zu Beginn der Testung auf das Vorhandensein von Stereotypen in der Bevölkerung aufmerksam gemacht; es wird beispielsweise angegeben, dass Frauen bessere Leistungen bei Gedächtnisaufgaben erzielen. Die Instruktion der zweiten Bedingung weist darauf hin, dass es keine geschlechtsspezifischen Unterschiede gibt. In der dritten Versuchsbedingung werden die SchülerInnen über das Phänomen aufgeklärt, dass bestimmte Meinungen über Stereotype der Gesellschaft zur schlechteren Bearbeitung von Aufgaben führen können. Die vierte Untersuchungsbedingung weist die SchülerInnen ebenfalls auf Stereotypen in der Bevölkerung hin (siehe erste Bedingung), allerdings werden die Fähigkeiten der Männer den Frauen zugeschrieben und umgekehrt. In der Instruktion wird zum Beispiel diesmal angegeben, dass Männer bessere Leistungen bei Gedächtnisaufgaben erzielen.

Jede der genannten Stereotype Threat Bedingungen wird zusätzlich in drei verschiedene Untersuchungsbedingungen unterteilt, um das Risikoverhalten zu messen. In der ersten Bedingung zur Messung des Risikoverhaltens werden die SchülerInnen dazu animiert zu raten, wenn sie die richtige Lösung nicht wissen. Die zweite Bedingung enthält keine Instruktion zum Risikoverhalten. In der dritten Bedingung werden die SchülerInnen darauf hingewiesen keine Items auszulassen.

Kognitive Aufgaben:

Es werden EMS-ähnliche Aufgaben jener Subtests („Quantitative und formale Probleme“, „Schlauchfiguren“, „Textverständnis“, „Med.-naturwissenschaftliches Grundverständnis“, „Diagramme und Tabellen“ und „Figuren lernen“ und „Fakten

lernen“) aus dem EMS vorgegeben, die 2008 signifikante Unterschiede zwischen den Geschlechtern gezeigt haben.

Zur Erfassung der Theory of Mind werden 3 selbstkonstruierte Geschichten vorgelegt.

Es werden Fragebögen zur Erfassung der Selbstwirksamkeit und zur Leistungsmotivation drei kurze Fragebögen vorgelegt. Des Weiteren werden soziodemografische Variablen und Interessenschwerpunkte, Wunsch nach Medizinstudium, Ausbildung (AHS-Zweig), Beschäftigung mit EMS-Vorbereitung usw. erfasst.

Stichprobe:

Es sollen etwa 1200 SchülerInnen der 8. Klassen von AHS (n=50 pro Zelle) - gestreut über verschiedene Standorte und Ausbildungszweige (u.a. naturwissenschaftlich, humanistisch, neusprachlich) in Wien rekrutiert werden.

Die gesamte Testung wird 2 Unterrichtsstunden umfassen und in Gruppentestungen durchgeführt.

A3. Einverständniserklärung

Probandeninformation und Einwilligungserklärung
zur Teilnahme an der Studie

Genderaspekte (Geschlechtsabhängige Aspekte) im Bezug auf den Eignungstest im
Medizinstudium (EMS-Test)

Sehr geehrte Studienteilnehmerin, sehr geehrter Studienteilnehmer!

Wir laden Sie ein an der oben genannten wissenschaftlichen Studie teilzunehmen. Die Aufklärung darüber erfolgt in einem ausführlichen Gespräch vor dem Test.

Ihre Teilnahme an dieser wissenschaftlichen Studie erfolgt freiwillig. Sie können jederzeit ohne Angabe von Gründen aus der Studie ausscheiden. Die Ablehnung der Teilnahme oder ein vorzeitiges Ausscheiden aus dieser Studie hat keine nachteiligen Folgen.

Wissenschaftliche Studien sind notwendig, um verlässliche neue Forschungsergebnisse zu gewinnen. Unverzichtbare Voraussetzung für die Durchführung einer wissenschaftlichen Studie ist jedoch, dass Sie Ihr Einverständnis zur Teilnahme an dieser Studie schriftlich erklären. Bitte lesen Sie den folgenden Text sorgfältig durch und zögern Sie nicht Fragen zu stellen.

Bitte unterschreiben Sie die Einwilligungserklärung nur

- wenn Sie Art und Ablauf der wissenschaftlichen Studie vollständig verstanden haben,
- wenn Sie bereit sind, der Teilnahme zuzustimmen und
- wenn Sie sich über Ihre Rechte als Teilnehmer an dieser wissenschaftlichen Studie im Klaren sind.

Zu dieser wissenschaftlichen Studie, sowie zur Probandeninformation und Einwilligungserklärung wurde von der zuständigen Ethikkommission eine befürwortende Stellungnahme abgegeben.

1. Zweck der Studie

Ziel dieser Studie ist es, den Einfluss sozialer und biologischer Faktoren auf Leistungsfähigkeiten während Testsituationen zu untersuchen und Unterschiede zwischen männlichen und weiblichen Jugendlichen zu beurteilen. Die Ergebnisse sollen im Weiteren helfen, Eignungstests für beide Geschlechter fair zu gestalten.

2. Teilnahme

Diese wissenschaftliche Studie wird an der Medizinischen Universität Wien durchgeführt, es werden insgesamt **1200 SchülerInnen** aus allgemeinbildenden höheren Schulen (AHS) daran teilnehmen.

Die Teilnahme an dieser wissenschaftlichen Studie wird voraussichtlich 2 Unterrichtsstunden dauern.

Folgende Maßnahmen werden ausschließlich aus Studiengründen durchgeführt:

Im Rahmen der Studie werden Sie gebeten, einen EMS (Eignungstest für Medizinstudium) ähnlichen Test mit Fragen, in denen sich im Jahr 2008 Unterschiede zwischen den Geschlechtern gezeigt haben, zu beantworten. Im Rahmen der Befragung werden Ihnen spezielle Anleitungen gegeben.

Am Ende des Tests sind kurze Fragebögen zur Erfassung der Selbstwirksamkeit und zur Leistungsmotivation auszufüllen. Außerdem werden Ihnen jeweils vor und nach dem Test eine Speichelprobe zur Untersuchung verschiedener geschlechtsspezifischer und belastungsabhängiger Botenstoffe (Hormone), wie Cortisol, Testosteron, Östradiol, Progesteron, DHEA, Dopamin und Chromogranin A, entnommen.

3. Nutzen einer Teilnahme an der wissenschaftlichen Studie

Sie haben durch die Teilnahme an dieser wissenschaftlichen Studie die Möglichkeit, Einblicke in Ihre eigenen Fähigkeiten der Selbstkompetenz und Leistung in Stresssituationen zu gewinnen und Ihre Schwächen durch gezieltes Verhalten richtig zu beeinflussen.

4. Risiken, Beschwerden und Begleiterscheinungen

Die Teilnahme an dieser Studie ist mit keinen Risiken verbunden.

5. Verpflichtungen

Diese Studie hat keinerlei Auswirkungen auf die Lebensführung und ist mit keinen weiteren Verpflichtungen verbunden.

Sie können die Teilnahme an der Studie jederzeit und ohne Angabe besonderer Gründe von sich aus abbrechen, wobei keine Nachteile erwachsen.

Der Versuchsleiter ist berechtigt die Untersuchung zu jeder Zeit von sich aus abzubrechen, wenn er es für erforderlich hält.

6. In welcher Weise werden die im Rahmen dieser Studie gesammelten Daten verwendet?

Es haben nur die Studienleiter/innen und deren Mitarbeiter/innen Zugang zu den vertraulichen Daten. Weiters können Beauftragte von in- und ausländischen Gesundheitsbehörden sowie der zuständigen Ethikkommission Einsicht in diese Daten nehmen, um die Richtigkeit der Aufzeichnungen zu überprüfen. Diese Personen unterliegen einer gesetzlichen Verschwiegenheitspflicht.

Die Weitergabe der Daten im In- und Ausland erfolgt ausschließlich zu statistischen Zwecken in verschlüsselter (nur „indirekt personenbezogener“) oder anonymisierter Form, das heißt, Sie werden nicht namentlich genannt. Auch in etwaigen Veröffentlichungen der Daten dieser Studie werden Sie nicht namentlich genannt.

Die Studienleiter/innen und ihre Mitarbeiter/innen unterliegen im Umgang mit den Daten den Bestimmungen des österreichischen Datenschutzgesetzes 2000 in der jeweils geltenden Fassung.

Wenn Sie Ihre Einwilligung zurückziehen und damit Ihre Teilnahme vorzeitig beenden, werden keine neuen Daten mehr über Sie erhoben. Auf Grund gesetzlicher Dokumentationspflichten kann jedoch weiterhin für einen gesetzlich festgelegten Zeitraum eine Einsichtnahme in Ihre personenbezogenen Daten zu Prüfzwecken durch autorisierte, zur Verschwiegenheit verpflichtete Personen erfolgen.

7. Entstehen für die Teilnehmer Kosten? Gibt es einen Kostenersatz oder eine Vergütung?

Durch Ihre Teilnahme an dieser Studie entstehen für Sie keine zusätzlichen Kosten. Ein Kostenersatz bzw. eine Vergütung ist nicht vorgesehen.

8. Möglichkeit zur Diskussion weiterer Fragen

Für weitere Fragen im Zusammenhang mit dieser wissenschaftlichen Studie steht Ihnen die Studienleitung gerne zur Verfügung. Auch Fragen, die die Rechte als Teilnehmer an dieser Studie betreffen, werden Ihnen gerne beantwortet.

Name der Kontaktpersonen: Univ.-Prof. Dr. Alexandra Kautzky-Willer

Ständig erreichbar unter: +43 1 40400 2126

Name der Kontaktperson: a.o. Univ-Prof. Mag. Dr. Ulrike Willinger

Ständig erreichbar unter: +43 1 40400 3106

14. Einwilligungserklärung

Name des Probanden in Druckbuchstaben:
.....

Ich erkläre mich bereit, an der Studie „Gender und EMS-Test“ teilzunehmen.

Ich bin von Herrn/Frau ausführlich und verständlich über mögliche Belastungen und Risiken, sowie über Wesen, Bedeutung und Tragweite der Untersuchung sowie die sich für mich daraus ergebenden Anforderungen aufgeklärt worden. Ich habe darüber hinaus den Text dieser Probandenaufklärung und Einwilligungserklärung, die insgesamt 4 Seiten umfasst gelesen. Aufgetretene Fragen wurden mir von der Studienleitung verständlich und genügend beantwortet. Ich hatte ausreichend Zeit, mich zu entscheiden. Ich habe zurzeit keine weiteren Fragen mehr.

Ich werde den Anordnungen, die für die Durchführung der Studie erforderlich sind, Folge leisten, behalte mir jedoch das Recht vor, meine freiwillige Mitwirkung jederzeit zu beenden.

Ich bin zugleich damit einverstanden, dass meine im Rahmen dieser Studie ermittelten Daten gespeichert werden. Mir ist bekannt, dass zur Überprüfung der Richtigkeit der Datenaufzeichnung Beauftragte der zuständigen Behörden und der Ethikkommission bei der Studienleitung Einblick in meine personenbezogenen Daten nehmen dürfen.

Beim Umgang mit den Daten werden die Bestimmungen des Datenschutzgesetzes 2000 beachtet.

Eine Kopie dieser Probandeninformation und Einwilligungserklärung werde ich auf Wunsch erhalten und das Original verbleibt bei der Studienleitung.

.....

(Datum und Unterschrift des Probanden)

.....

(Datum und Unterschrift eines Elternteils, falls Proband nicht volljährig ist)

.....

(Datum, Name und Unterschrift des verantwortlichen Studienleiters)

(Der Proband erhält auf Wunsch eine unterschriebene Kopie der Probandeninformation und Einwilligungserklärung, das Original verbleibt im Studienordner des Studienleitung.)

A4. Lebenslauf

Zur Person

Name: Katrin Anzirk

Geburtsdatum: 27.05.1983

Geburtsort: Wien

Staatsbürgerschaft: Österreich

E-mail: a0251713@unet.univie.ac.at

Ausbildung

- | | |
|--------------------------|---|
| 1994-2002 | Dreisprachiges, humanistisches Gymnasium Wien 14 (Englisch, Französisch, Italienisch), |
| 2003 (WS) | Psychologiestudium an der Universität Wien. |
| 10/2003 – 02/2004 | Lehrgang für Werbung und Marktkommunikation am WIFI St. Pölten mit sehr gutem Erfolg absolviert. |

Berufliche Tätigkeiten

- | | |
|-------------------------|---|
| 2/2005 – 3/2007 | Seniorenhilfe Junge Panther: Betreuung älterer Menschen |
| 8/2007 - 8/2008 | Wiener Assistenzgenossenschaft: Tätigkeit als persönliche Assistentin |
| 9/2008 - 2/2010 | GfK-Austria: Marktforschung |
| 9/2010 - 12/2010 | Praktikum in der Gerontopsychiatrischen Tagesklinik 19/3 im Sozialmedizinischen Zentrum Baumgartner Höhe |
| Seit März 2010 | Verein LOK-Leben ohne Krankenhaus: Tätigkeit als Betreuerin in zwei Wohngemeinschaften |