



universität
wien

Diplomarbeit

Titel der Arbeit

*Geschlechtsspezifische Unterschiede in Bezug auf
Gedächtnisleistungen im Rahmen EMS-ähnlicher
Aufgaben unter Berücksichtigung von Stereotype
Threat und Risikoverhalten*

Verfasserin

Helene Antonia Lager

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im Oktober 2011

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: Univ.- Prof. Mag. Dr. Ulrike Willinger

Danksagung

An dieser Stelle möchte ich mich bei allen Menschen bedanken, die mich in der Zeit meiner Diplomarbeit begleitet und unterstützt haben.

Zu allererst ein großes Dankeschön meiner Betreuerin Ao. Univ.-Prof. Mag. Dr. Ulrike Willinger für die wertvolle Unterstützung und hilfreichen Anregungen während der Erstellung dieser Arbeit. Vielen Dank für die Herausforderung durch dieses hochaktuelle und spannende Thema.

Ein Dankeschön gilt den teilnehmenden Schulen in Wien, Niederösterreich und dem Burgenland, durch deren Bereitschaft und Kooperation dieses Projekt erst in die Realität umgesetzt werden konnte.

Bei meiner Familie, insbesondere meiner Großmutter und meinem Onkel für die Unterstützung aus der Ferne während meiner gesamten Studienzeit, wie ich sie mir liebevoller und großzügiger nicht hätte vorstellen können.

Meiner Mutter, für ihre emotionale Unterstützung, für die großartige Motivationsarbeit und dass sie mir immer Mut gemacht hat an sich selbst zu glauben. Ohne die Gewissheit, dass sie immer hinter mir steht, hätte ich mein Studium nicht absolvieren können und dafür möchte ich mich aus tiefstem Herzen bedanken.

Spezieller Dank gilt an dieser Stelle meinen Kolleginnen Katrin Anzirk, Gabriele Hangl, Silvia Hameseder und Anita Teufl für die tolle Zusammenarbeit und die vielen Stunden, die wir gemeinsam an dieser Forschungsarbeit verbringen durften. Durch Euch wird diese Zeit für mich für immer unvergessen bleiben.

Meinem Großvater, dass er immer an meiner Seite ist und mich auf jede erdenkliche Art und Weise gefördert und unterstützt hat. Du bist mein größtes Vorbild, dir seien diese Zeilen gewidmet.

INHALTSVERZEICHNIS

1	EINLEITUNG	9
2	THEORIETEIL	11
2.1	Fakten zum Eignungstest für das Medizinstudium (EMS)	11
2.1.1	Kritische Auseinandersetzung mit dem EMS- Testmaterial.....	13
2.2	Lernen und Gedächtnis	17
2.2.1	Definition von Lernen und Gedächtnis	17
2.2.2	Gedächtnissysteme.....	19
2.2.3	Mehrspeichermodelle	19
2.2.4	Sensorisches Gedächtnis	20
2.2.5	Kurzzeitgedächtnis.....	21
2.2.5.1	Modell des Arbeitsgedächtnis nach Baddeley	21
2.2.6	Langzeitgedächtnis	22
2.2.7	Phänomene welche den Erwerb von Lerninhalten beeinflussen	23
2.2.7.1	Positionseffekte: Primacy- und Recency- Effekt	24
2.2.7.2	Kodierungsspezifität und Hinweisreize beim Abruf von Information	25
2.2.8	Geschlechtsunterschiede im Zusammenhang mit Lern- und Merkfähigkeit	26
2.3	Stereotype und Stereotype Threat.....	31
2.3.1	Definition von Stereotype.....	32
2.3.2	Bedrohung durch Stereotype- Stereotype Threat	33
2.3.2.1	Stereotype Threat und Geschlecht	34
2.3.2.2	Stereotype Threat und Ängstlichkeit	36
2.3.2.3	Stereotype Threat und Entscheidungsverhalten.....	36
2.3.2.4	Stereotype Threat und Gedächtnis	37
2.3.2.5	Reduzierung von Stereotype Threat	39
2.4	Risikoverhalten	43
2.4.1	Definition von Risikoverhalten	43
2.4.1.1	Definition von Entscheidung	44
2.4.1.2	Entscheidung unter Risiko	45
2.4.1.3	SEU-Theorie- Subjectively Expected Utility Modell.....	46

2.4.2	Risikoverhalten und Geschlecht	47
2.4.2.1	Geschlechterunterschiede im Risikoverhalten im Leistungsbereich mit Multiple Choice Antwortformat	49
2.4.3	Game of Dice Task (GDT)	50
3	EMPIRISCHER TEIL.....	53
3.1	Zielsetzung, Fragestellung und Hypothesen	53
3.2	Methode.....	53
3.2.1	Untersuchungsplan und intendierte Stichprobe	54
3.2.2	Erhebungsinstrumente	56
3.2.2.1	Experimentelle Versuchsbedingungen.....	57
3.2.2.2	Soziodemographischer Fragebogen	58
3.2.2.3	Wissentest.....	58
3.2.2.4	EMS-ähnliche Aufgaben	60
3.2.2.5	Theory of mind	64
3.2.2.6	Leistungsmotivation	65
3.2.2.7	Selbstwirksamkeitserwartung	65
3.2.2.8	Game of Dice Task (GDT).....	66
3.3	Untersuchung	67
3.3.1	Untersuchungsdurchführung.....	67
3.3.2	Beschreibung der Stichprobe.....	68
3.3.2.1	Nationalität der SchülerInnen	70
3.3.2.2	Lebensalter der Gesamtstichprobe	70
3.3.2.3	Schulzweige der SchülerInnen	72
3.3.2.4	Wiederholen einer Schulstufe	74
3.3.2.5	Wunsch nach einem Medizinstudium und Vorbereitung auf EMS	77
3.3.2.6	Verteilung der Geschlechter in die Versuchsbedingungen	84
3.3.2.7	Deskriptivstatistik der EMS – ähnlichen Aufgaben	90
3.3.2.8	Relative Lösungshäufigkeiten in Abhängigkeit der Stichprobe	91
3.4	Auswertung und Ergebnisse	92
3.4.1	Auswertungsverfahren	93
3.4.2	Überprüfung der experimentellen Variationen mittels Wissenstest.....	94
3.4.3	Reliabilitätsanalysen	102

3.4.3.1	Reliabilitätsanalysen der EMS-ähnlichen Aufgaben und Game of Dice Task (GDT) mittels Cronbach's Alpha	102
3.4.3.2	Reliabilitätsanalysen der Faktorstufenkombinationen Stereotype Threat und Risikoverhalten der 6 EMS- ähnlichen Untertests.....	104
3.4.3.3	Reliabilitätsanalyse des EMS- ähnlichen Gesamttest und der Untertests zur Erfassung der Merkfähigkeit	107
3.4.4	Interkorrelationen der EMS- ähnlichen Untertests	108
3.4.5	Darstellung der Hypothesenbezogenen Ergebnisse	109
3.4.5.1	Darstellung der univariaten Ergebnisse des EMS- ähnlichen Untertests: Figuren Lernen und Fakten Lernen	110
3.4.5.2	Nebenhypothese der EMS- ähnlichen Untertests in Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den EMS	117
3.4.5.3	Hypothesenbezogenen Ergebnisse des Game of Dice Task (GDT)	124
4	DISKUSSION UND AUSBLICK	128
5	ZUSAMMENFASSUNG DER ERGEBNISSE.....	135
6	ABSTRACT	138
7	LITERATURVERZEICHNIS	140
8	ABBILDUNGSVERZEICHNIS	147
9	TABELLENVERZEICHNIS.....	148
10	ANHANG	151

Anmerkung

Die im Rahmen dieser Diplomarbeit durchgeführte empirische Untersuchung wird auf Basis einer gemeinsamen Datenerhebung mit den Kolleginnen Katrin Anzirk, Gabriele Hangl, Silvia Hameseder und Anita Teufl dargestellt. Die daraus resultierenden Arbeiten basieren auf derselben Stichprobe und einem vergleichbaren Erkenntnisstand. Aus diesem Grund kann es Überschneidungsbereiche geben, die nicht als Plagiat zu werten sind. In diesem Sinne behandeln all diese Arbeiten jeweils einen spezifischen Themenschwerpunkt und sind als eigenständige Beiträge zur Forschungsfrage zu betrachten.

1 EINLEITUNG

Die vorliegende Arbeit ist der Teil eines größer angelegten Forschungsprojektes und geht anhand verschiedener Aspekte der Frage nach, warum Frauen im Eingangstest für das Medizinstudium (EMS) bedeutend schlechter abschneiden als ihre männlichen Kollegen, obwohl sich diese geschlechtsspezifischen Leistungsunterschiede während des Medizinstudiums nahezu auflösen und in Folge sogar mehr Frauen ein Medizinstudium abschließen als Männer (Spiel, Schober & Litzenberger, 2008).

Hauptziel ist es, Geschlechterunterschiede und Einflussfaktoren in Bezug auf den Eignungstest für das Medizinstudium (EMS) zu untersuchen, mit besonderer Betonung auf etwaige Unterschiede von Männern und Frauen hinsichtlich der Leistungen, welche die Merkfähigkeit betreffen.

In der Theorie wird eine Reihe möglicher Ursachen beschrieben, um Geschlechterunterschiede in Bezug auf kognitive Leistungsfähigkeit zu erklären. Als einer der wesentlichen Faktoren in Zusammenhang mit Geschlechterunterschieden in Leistungssituationen, auf welche in der vorliegenden Diplomarbeit eingegangen wird, gelten unter anderem die sozialen Phänomene *Stereotype Threat* und *Risikobereitschaft*.

In Untersuchungen konnte oftmals bestätigt werden, dass die Aktivierung von *Stereotype Threat* („Frauen sind schlecht in Mathematik“) die Testleistung von Personen, insbesondere von Frauen, insofern verringert, als es ihrem eigentlichen Leistungsvermögen entsprechen würde (vgl. Keller & Dauenheimer, 2003; Spencer, Steele & Quinn, 1999). Ebenso konnten Geschlechtsunterschiede in Bezug auf das Phänomen der *Risikobereitschaft*, welche zumeist mit erhöhter *Risikobereitschaft* zugunsten der Männer ausfielen, weitgehend bestätigt werden (vgl. Eckel & Grossmann, 2008).

Während die Kolleginnen Katrin Anzirk, Gabriele Hangl, Silvia Hameseder und Anita Teufl sich unter anderem mit Geschlechtsunterschieden im EMS

in Bezug auf Theory of Mind, Selbstwirksamkeitserwartung, Leistungsmotivation und schlussfolgerndem Denken beschäftigen, konzentriert sich diese Diplomarbeit schwerpunktmäßig auf die Unterschiede der Geschlechter im Bereich der Merkfähigkeit, die im Eignungstest für das Medizinstudium (EMS) in Form von zwei Untertests (*Fakten-* und *Figuren Lernen*) erfasst werden.

Diese Arbeit soll einen Beitrag darstellen, um die Rolle des Konstrukts Merkfähigkeit in Anbetracht der beobachtbaren Leistungsunterschiede im Eignungstest für das Medizinstudium (EMS) zu klären.

Im theoretischen Teil sollen zuerst grundlegende Informationen zum Eignungstest im Medizinstudium (EMS) gegeben werden, bevor des Weiteren auf die Merkfähigkeit und das Lernen anhand von Theorien und in weiterer Folge von Studien, welche Geschlechterunterschiede im Bereich der Merkfähigkeit untersuchen, eingegangen werden. In einem weiteren Teil wird das Phänomen von *Stereotype Threat* erklärt und aktuelle Forschungsarbeiten, unter anderem auch hinsichtlich seines Einflusses auf die Merkfähigkeit diskutiert.

Im darauffolgenden Kapitel wird das Konstrukt Risikoverhalten theoretisch anhand von Theorien und Studien, die wiederum in Bezug auf etwaige Unterschiede zwischen Männern und Frauen abzielen, erklärt.

Im empirischen Teil wird der wechselseitige Einfluss der im Theorieteil beschriebenen sozialen Phänomene Stereotype Threat und Risikoverhalten auf die kognitiven Leistungen im Medizineingangstest (EMS), insbesondere in Bezug auf mögliche geschlechtsspezifische Unterschiede, geprüft. Darüber hinaus wird im empirischen Teil dieser Arbeit ein Überblick über die Rahmenbedingungen und die verwendeten Erhebungsinstrumente gegeben, bevor im Anschluss daran eine detaillierte Beschreibung der untersuchten Stichprobe erfolgt. Schließlich werden wesentliche Ergebnisse, wieder mit dem Schwerpunkt der EMS-ähnlichen Untertests zur Erfassung der Merkfähigkeit, dargestellt.

Im Diskussionsteil werden die Zusammenhänge zwischen den im Theorieteil beschriebenen Studien und der aktuell durchgeführten Untersuchung hergestellt sowie Perspektiven und Anhaltspunkte für eine faire Gestaltung des Eingangstest für das Medizinstudium (EMS) erörtert.

2 THEORIETEIL

Der Theoretische Teil dieser Arbeit soll einen Überblick über die Konstrukte geben, welche im Rahmen des empirischen Teils der vorliegenden Arbeit von besonderem Interesse sind.

2.1 FAKTEN ZUM EIGNUNGSTEST FÜR DAS MEDIZINSTUDIUM (EMS)

Aufgrund der Entscheidung des Europäischen Gerichtshofs hinsichtlich des Zugangs zu Hochschulen in Österreich und der Erneuerung des Universitätsgesetzes 2002, kam es zur Einführung von Zugangsbeschränkungen an Universitäten Österreichs. Seit dem Studienjahr 2006/07 gelangen an den Universitäten Wien, Innsbruck und Graz in einigen Studienrichtungen Auswahlverfahren zum Einsatz (Spiel et al., 2008).

Allgemeines Ziel von Aufnahmeverfahren an Universitäten ist, die Fähigkeit zu studieren und die Vorhersage des Studienerfolgs zu erfassen (Spiel, Litzenberger & Haiden, 2007).

Das Auswahlverfahren für die Eignung zum Medizinstudium an den Universitäten Wien und Innsbruck wird seither mittels eines Schweizer Testverfahren (EMS) ermittelt (Spiel et al., 2008). Durch die Vorgabe des Eignungstests EMS wird seither die Studienplatzvergabe für alle angehenden Studierenden an den Universitäten Wien und Innsbruck ermittelt.

Der Eignungstest für das Medizin Studium (EMS) baut auf dem deutschen Test für Medizinische Studiengänge (TMS) auf und wurde in der Schweiz weiterentwickelt, wo er seit 1998 zum Einsatz kommt und insgesamt aus 10 Untertests besteht (Mallinger et al., 2009).

Für Graz wurde ein eigener Test zur Eignung für das Medizinstudium entwickelt, der verschiedene medizinrelevante Bereiche, wie beispielsweise *Biologie, Chemie, Physik, Mathematik und Textverständnis* mit insgesamt 200 Aufgaben beinhaltet (Spiel et al., 2008).

Grundsätzlich schneiden Frauen in beiden Eignungstests zum Medizinstudium etwas schlechter ab als ihre männlichen Kollegen und BewerberInnen aus Österreich etwas schlechter als BewerberInnen aus Deutschland. Diese Tatsache hatte deutliche Auswirkungen auf die Geschlechts- und Nationalverteilung derjenigen, die zum Medizinstudium zugelassen wurden. In diesem Zusammenhang wurde schließlich eine sogenannte Quotenregelung für die Aufnahme zum Studium der Medizin (75% ÖsterreicherInnen, 20% EU-BürgerInnen, 5% Nicht-EU-BürgerInnen) eingeführt, um die Aufnahme der deutschen BewerberInnen, die etwas bessere Leistungen im Eignungstest als ihre Österreichischen KollegInnen zeigen, zu reglementieren (Spiel et al., 2008).

Tabelle 1: Untertests im EMS-Aufnahmeverfahren (modifiziert nach Hänsgen & Spicher, 2007)

EMS-Untertests	Geprüfte Fähigkeiten
Muster ordnen	Differenzierte visuelle Wahrnehmung
Medizinisch-naturwissenschaftliches Grundverständnis	Verständnis für Fragen der Medizin und der Naturwissenschaften
Schlauchfiguren	Räumliches Vorstellungsvermögen
Quantitative und formale Probleme	Quantitatives Problemlösen im medizinisch-naturwissenschaftlichen Kontext
Textverständnis	Verarbeitung von komplexem Textmaterial
Figuren lernen	Behalten von figuralem Material
Fakten lernen	Behalten von verbalem Material
Planen und Organisieren	Effiziente Selbstorganisation im Studium
Diagramme und Tabellen	Analyse und Interpretation von Diagrammen und Tabellen
Konzentriertes und sorgfältiges Arbeiten	rasches und konzentriertes Arbeiten

2.1.1 KRITISCHE AUSEINANDERSETZUNG MIT DEM EMS- TESTMATERIAL

Ziel der Untersuchung von Spiel, Schober und Litzenberger (2008) war die Evaluation der Tests für die Eignung zum Medizinstudium in Österreich mit besonderer Berücksichtigung der Fairness hinsichtlich des Geschlechts, der Nationalität, der sozialen Herkunft und des Schultyp der BewerberInnen.

Besonderes Augenmerk galt der Analyse dreier Bereiche:

- 1) Identifikation jener Faktoren, welche einen wesentlichen Einfluss auf die Leistung im Test zur Eignung im Medizin Studium haben,
- 2) Fairness der Verfahren, die zur Auswahl zu einem Medizinstudium herangezogen werden,

3) Bildungsweg und Erfolg im Test

Für die Klärung der Fragestellung wurden die notwendigen Daten im Rahmen des Auswahlverfahren 2007/08 von allen BewerberInnen erhoben. Mit einem neu entwickelten Fragebogen wurden Information bezüglich familiärem Hintergrund, schulischem Werdegang und den Erfahrungen mit Eignungstests im Rahmen der Auswahl zum Medizinstudium erfasst. Des Weiteren wurden Selbsteinschätzungen, welche die Situation während des Tests betrafen und die Wichtigkeit ein Medizinstudium zu absolvieren ermittelt (Spiel et al., 2008).

Die Ergebnisse weisen darauf hin, dass Unterschiede in der Testleistung unter anderem mit den Schulnoten zusammenhängen. Jene BewerberInnen, die den Eignungstest erfolgreich absolviert haben, verfügen im Gegensatz zu ihren nicht erfolgreichen KollegInnen über deutlich bessere Schulnoten und zeichnen sich durch mehr Sicherheit hinsichtlich ihrer Fähigkeiten in mathematisch-naturwissenschaftlichen Themenbereichen als ihre nicht erfolgreichen MitbewerberInnen aus. Zudem konnte gezeigt werden, dass jene Personen, welche den Eignungstest erfolgreich absolvierten, öfter zu dieser Art von Auswahlverfahren angetreten sind und sie bezüglich eines erfolgreichen Testergebnisses zuversichtlicher als nicht erfolgreiche sind. Bezüglich der unterschiedlichen Nationalitäten, die zum Eignungstest für das Medizinstudium angetreten waren, konnte festgestellt werden, dass deutsche BewerberInnen insgesamt häufiger zu einem Studieneignungstest angetreten waren und ihre Eltern im Mittel eine höhere Schulausbildung aufwiesen als die der österreichischen BewerberInnen (Spiel et al., 2008).

In Bezug auf die Fairness des Auswahlverfahrens zum Medizinstudium ergab sich anhand der Analysen für den EMS ein insgesamt bedenkliches Resultat. Die Einzelaufgaben wurden in Bezug auf ihre Trennschärfen und auf ihre Schwierigkeit hin untersucht. Keine der untersuchten Aufgaben zeigt eine gute Trennschärfe und keiner der Untertests verfügt über eine ausreichende Genauigkeit (Reliabilität). Des Weiteren zeigten sich

Probleme der Verrechnungsfairness, d.h. das Lösen der jeweiligen Aufgabe im Eignungstest ist nicht nur vom Wissen der Person abhängig, sondern auch von anderen Faktoren, wie beispielsweise dem Geschlecht. Es kann davon ausgegangen werden, dass einige der Aufgaben bei Männern und Frauen bzw. bei österreichischen und deutschen StudienbewerberInnen nicht exakt dasselbe erfassen (Spiel et al., 2008).

Frauen und Männer unterscheiden sich in Bezug auf ihre Leistungen im Eignungstest systematisch voneinander, Frauen zeigen in sieben (*Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Planen und Organisieren, Textverständnis, Konzentriertes und Sorgfältiges Arbeiten, Quantitative und Formale Probleme und Diagramme und Tabellen*) von zehn Untertests schlechtere Leistungen als Männer. Allerdings soll an dieser Stelle erwähnt werden, dass sich im Laufe des Medizinstudiums jene anfänglichen Leistungsunterschiede zur Gänze aufheben und letztlich mehr Frauen das Medizinstudium abschließen (Spiel et al., 2008).

Spiel, Schober und Litzenberger (2008) gehen davon aus, dass der Ursprung der unterschiedlichen Leistungen möglicherweise in unterschiedlichen Sozialisationsprozessen begründet werden kann. Unterschiede im Geschlecht, die noch im Kindergartenalter bzw. in der Grundschule nicht feststellbar sind, scheinen mit zunehmendem Alter vermehrt aufzutreten, vor allem in mathematisch-naturwissenschaftlichen Schulfächern. Von heranwachsenden Buben werden vor allem in diesen Bereichen bessere Leistungen erwartet, gute Leistungen der Mädchen als weniger wichtig für den späteren Beruf erachtet und demnach weniger gefördert. Die Ergebnisse zeigen deutlich, dass Frauen, obwohl sie gute Schulnoten vorzeigen können, schlechter im Eignungstest abschneiden als ihre männlichen Mitbewerber. Aufgrund dieser Ergebnisse wurden die Schulnoten und das Abschneiden im Eignungstest getrennt für die Geschlechter ermittelt, wonach Frauen mit den gleichen Schulnoten deutlich schlechter im Test abschneiden. Aufgrund der Tatsache des besseren Abschneidens im Eignungstest der deutschen Studien-

bewerberInnen bei gleichzeitig schlechteren Schulnoten, wurde zusätzlich die Nationalität analysiert. Die Ergebnisse zeigten, dass weibliche Bewerber aus Österreich die deutlich schlechtesten Leistungen im Eignungstest erbrachten. Jedoch hat die Interpretation dieses Ergebnisses aufgrund der geringen Anzahl der BewerberInnen aus Deutschland im Vergleich zu österreichischen nur eingeschränkte Aussagekraft (Spiel et al., 2008). Hinsichtlich der Entwicklung der Aufgaben im Rahmen der Vorgabe des Eignungsverfahrens zum Medizinstudium (EMS) wurde darauf geachtet, dass kein spezifisches Vorwissen zur Bearbeitung der Aufgaben notwendig ist und somit die Trainierbarkeit des EMS möglichst gering bleibt (Mallinger et al., 2009).

In einer prospektiven Untersuchung von Frischenschlager, Mitterauer und Haidinger (2005) wurden anhand von StudentInnen, die sich am Beginn ihres Medizinstudiums an der Universität Wien befanden, Indikatoren für die Leistung (Schulnoten, fachspezifische Vorkenntnisse, allgemeine und fachspezifische Lernleistung) ermittelt, um sie im Anschluss daran mit dem Studienerfolg in Zusammenhang zu bringen. Anliegen der Untersuchung war, in welcher Weise Faktoren der Leistung mit dem Studienerfolg in Beziehung zu setzen sind und inwieweit sich diese Leistungsfaktoren als Auswahlkriterien für die Aufnahme zum Medizinstudium eignen (Frischenschlager et al., 2005). Insgesamt wurden 839 StudienanfängerInnen der Medizinischen Universität Wien in den ersten Tagen des Studienjahres 2003/04 untersucht, um die prognostische Güte von Schulnoten, allgemeiner und fachspezifischer Lernfähigkeit und fachspezifische Vorkenntnisse zu ermitteln. Als Merkmal für den Studienerfolg wurde das Bestehen der großen Abschlussprüfung (SIP) am Ende des ersten Studienjahres ausgewählt. Neben den Schulnoten in Physik, Chemie, Mathematik und Englisch wurde die Dauer des Lateinunterrichts in Jahren erhoben und ob eine Schulklasse wiederholt wurde. Mittels 20 Fragen mit den Schwerpunkten Physik, Chemie und Biologie wurden die fachspezifischen Vorkenntnisse erfragt. Die allgemeine Lernleistung wurde mit einem Lern- und Merkfähigkeitstest

ermittelt. Die zweite Komponente der Lernleistung, nämlich die fachspezifische Lernfähigkeit und die Fähigkeit des ständigen Mitlernens, wie sie auch im Medizinstudium gefordert werden, konnte mit einer Prüfung im Umfang von fünfzehn Fragen über den Stoff einer Biochemie-Vorlesung, welche die StudentInnen zwei Tage zuvor besucht hatten, erfasst werden. Die Ergebnisse zeigen deutlich die hohe Vorhersehbarkeit von Schulleistungen. Wenn jedoch schulische Leistungen als Kriterium für die Aufnahme zu einem Medizinstudium herangezogen werden sollen, gilt als Voraussetzung die Feststellung der Leistung am Ende des Schulbesuches in standardisierter Form, welches in Österreich aufgrund des sehr breiten Angebots an Schultypen und Schulzweigen nicht als gegeben betrachtet werden kann (Frischschläger et al., 2005). Entgegen der Erwartung konnte auch die allgemeine Lernleistung nicht als Prädiktor für den Studienerfolg herangezogen werden. Fachspezifische Vorkenntnisse konnten mit dem Studienerfolg in Zusammenhang gebracht werden, jedoch unzureichend, um als Kriterium für die Auswahl verwendet zu werden. Insgesamt konnte festgestellt werden, dass keiner der in dieser Untersuchung erhobenen Indikatoren für Leistung als Verfahren zur Auswahl zu einem Medizinstudium geeignet ist, insbesondere aufgrund des in Österreich vorhandenen unterschiedlichen Schulsystems (Frischenschlager et al., 2005).

2.2 LERNEN UND GEDÄCHTNIS

Im folgenden Abschnitt werden Erklärungsansätze für Lernen und Gedächtnis erörtert und jene Bereiche beschrieben, die für die Gedächtnisleistung wesentlich sind.

2.2.1 DEFINITION VON LERNEN UND GEDÄCHTNIS

Gedächtnis stellt jene Instanz dar, welche der Fähigkeit obliegt Informationen aufzunehmen, diese zu speichern und zu gegebenem Zeitpunkt wieder abzurufen bzw. zu reproduzieren (Schermer, 2006).

Eine bedeutende Voraussetzung, die Lernen erst ermöglicht, stellt die Fähigkeit des Behaltens dar (Seel, 2003). Aufmerksamkeit, Lernen und Gedächtnis stehen miteinander in Verbindung und beeinflussen sich gegenseitig. Lernen bewirkt eine Verhaltensänderung, die durch Erfahrung bedingt wird, die Konsequenz von Erfahrung stellt das Gedächtnis dar und letztendlich werden Lern- und Gedächtnisprozesse durch Aufmerksamkeit erleichtert. Die Auseinandersetzung mit dem Themenbereich Gedächtnis ist gleichzeitig eine mit dem Bereich Lernen und umgekehrt (Lefrancois, 2006).

„Genauso wie das Lernen ist auch das Gedächtnis ein sog. hypothetisches Konstrukt, d.h. unserer direkten Beobachtung nicht zugänglich und muss aus seinen Effekten erschlossen werden.“ (Schermer, 2006, S.13)

Einige Funktionen, wenn auch in unterschiedlicher Schwerpunktsetzung, sind für die Leistung von Gedächtnis wesentlich (Schermer, 2006; Zimbardo & Gerrig, 2004).

Enkodierung: Zuallererst muss für die Enkodierung sichergestellt sein, dass die Aufmerksamkeit auf die Information gerichtet wird um in weiterer Folge in ein anderes Format verarbeitet zu werden. Damit ist jedoch nicht gemeint, dass eine vollkommen bewusste Aufmerksamkeitsleistung für die Verarbeitung erforderlich ist, denn Enkodierung ist auch bei Wahrnehmung der Information möglich, kann demnach auch unbewusst erfolgen.

Speicherung: Werden Informationen auf richtige Weise enkodiert, ermöglicht dies das Behalten der Informationen über eine bestimmte Zeitspanne.

Abruf: Informationen, die bereits im Gedächtnis gespeichert werden konnten, werden zu einem späteren Zeitpunkt wieder abgerufen.

Diese drei für das Gedächtnis wesentlichen Prozesse interagieren auf eine komplexe Art und Weise (Zimbardo & Gerrig, 2004).

Speicherung von Informationen, als auch Prozesse des Kodierens können nur in Folge von Ergebnissen des Abrufs erkannt werden, eine direkte Beobachtung dieser ist somit nicht möglich. Die Tätigkeit des Gedächtnisses kann durch erfolgreiche bzw. nicht erfolgreiche Abrufergebnisse bestimmt werden (Schermer, 2006).

2.2.2 GEDÄCHTNISSYSTEME

Im Folgenden sollen die verschiedenen Aspekte und unterschiedlichen Systeme von Gedächtnis beschrieben und erklärt werden.

2.2.3 MEHRSPEICHERMODELLE

Eine frühe Form der Mehrspeichermodelle stellten die sogenannten Zweispeichermodelle dar, bei denen es zu einer Unterscheidung in *Kurzzeitgedächtnis* (KZG) und *Langzeitgedächtnis* (LZG) kommt, die jeweils für kurz- bzw. langfristiges Behalten von Informationen zuständig sind. Den beiden Gedächtnisarten sind verschiedene Systeme zugehörig, der Kurzzeitspeicher und der Langzeitspeicher. Das Wiederholen bzw. das Memorieren (rehearsal) eines Lernstoffs ermöglicht den Übergang der Information vom Kurzzeitgedächtnis in das Langzeitgedächtnis (Schermer, 2006).

Schließlich wurde 1968 von Atkinson und Shiffrin das Zweispeichermodell mit der Unterscheidung in Kurz- bzw. Langzeitgedächtnis durch eine dritte Komponente, dem sogenannten *Sensorischen Gedächtnis*, welches dem Kurzzeitgedächtnis voraus geschaltet ist, erweitert (Atkinson & Shiffrin, 1968, zitiert nach Schermer, 2006, S. 117).

In folgenden Aspekten unterscheiden sich die drei Gedächtnissysteme voneinander (Schermer, 2006).

1. Kapazität: Die Fülle an Informationen, die verarbeitet werden können.

2. Haltezeit: Der Zeitraum in der die Informationen behalten werden kann.
3. Kodierung: Die Verschlüsselung der Informationen.

2.2.4 SENSORISCHES GEDÄCHTNIS

Das *sensorische* Register wird auch als Ultrakurzzeitgedächtnis bezeichnet und umfasst alle nicht bewussten und unmittelbaren Wirkungen von Reizeindrücken und wird in ein *ikonisches* und *echoisches* Gedächtnis unterteilt (Lefrancois, 2006).

Das *ikonische* Gedächtnis bezieht sich auf visuell zu verarbeitende Inhalte des sensorischen Gedächtnisses. Der *ikonische* Speicher ist in der Lage, sehr große Mengen von Informationen über einen Zeitraum von unter einer Sekunde zu speichern (Neisser, 1967). Das *echoische* Gedächtnis ist, ebenso wie das *ikonische*, in der Lage große Informationsmengen für eine kurze Zeit aufrechtzuerhalten, das *echoische* Gedächtnis bezieht sich jedoch auf auditive Reizmuster (Zimbardo & Gerrig, 2004).

In einer Untersuchung 1963 zum visuell *sensorischen* Gedächtnis von Sperling wurde VersuchsteilnehmerInnen mittels eines Tachistoskop drei Reihen mit je vier Buchstaben auf einem Bildschirm weniger als eine Sekunde gezeigt. Unmittelbar danach wurde den Versuchspersonen ein Ton dargeboten, den sie einer der drei Buchstabenreihen zuordnen konnten und somit aufgefordert wurden, genau diese Reihe zu reproduzieren. Die Ergebnisse zeigten, dass unter diesen Bedingungen ein Großteil der VersuchsteilnehmerInnen alle vier Buchstaben der Reihe wiedergeben konnten. Sollten sich die Personen an alle insgesamt 12 Buchstaben erinnern, so konnten sie im Mittel nur etwa 4,5 reproduzieren. Die experimentellen Untersuchungen von Sperling zeigen, dass eine kleine Anzahl von Reizmaterial über einen sehr kurzen Zeitraum nach der Präsentation zwar nicht bewusst ist, jedoch abgerufen und erinnert werden kann (Sperling, 1963, zitiert nach Lefrancois, 2006, S. 261).

2.2.5 KURZZEITGEDÄCHTNIS

Informationen aus dem sensorischen Gedächtnis werden ins Kurzzeitgedächtnis überführt und dort weiterverarbeitet um Informationen auch für einen etwas längeren Zeitraum, als es im sensorischen Register möglich ist, aufrechtzuerhalten (Schermer, 2006).

„Kurzzeitgedächtnis ist das, was zu einem gegebenen Zeitpunkt bewusst ist.“ (Lefrancois, 2006, S.262). Das Kurzzeitgedächtnis hat eine begrenzte Speicherkapazität, es ermöglicht, dass unmittelbar aufgenommene Informationen verfügbar bleiben und Informationen aus dem Langzeitgedächtnis abgerufen werden können. Es ist notwendig, Informationen im Kurzzeitgedächtnis zu wiederholen, ansonsten ist die Speicherung auch für dieses Gedächtnissystem begrenzt (Zimbardo & Gerrig, 2004). Die Speicherkapazität des Kurzzeitgedächtnisses kann mit der Gedächtnisspanne gleichgesetzt und durch sie ermittelt werden. Miller (1956) kam in seinen Untersuchungen zum Schluss, dass die Kapazität des Kurzzeitgedächtnisses bzw. die Gedächtnisspanne eines Erwachsenen auf 7 ± 2 Einheiten von Informationen festgelegt werden kann. Die Zusammenfassung von einzelnen Einheiten von Informationen in Hinblick auf ihre Bedeutung nennt man *chunking*, die kleinste Einheit von Information wird als *chunk* (Klumpen) bezeichnet und ermöglicht die Kapazität des Kurzzeitgedächtnisses zu vergrößern und sie ökonomischer zu nutzen (Schermer, 2006). Beispielsweise könnte man sich die Lebensmittel auf einer Einkaufsliste im Einzelnen folgendermaßen merken: Eier, Zucker, Mehl und Backpulver oder aber die Menge an *chunks*, in diesem Fall wären es vier, zu „*Großmutter's Kuchen*“ also zu einem *chunk* zusammenzufassen und letztlich die Gedächtnisspanne auszuweiten.

2.2.5.1 Modell des Arbeitsgedächtnis nach Baddeley

Das Arbeitsgedächtnis kann dem Begriff des Kurzzeitgedächtnisses gleichgesetzt werden und durch die Integration von Kurzzeitgedächtnis in den Kontext des Arbeitsgedächtnisses können diese Systeme weniger als Ort, sondern eher als ein Prozess verstanden werden (Zimbardo & Gerrig,

2004). Baddeley (1986, zitiert nach Zimbardo & Gerrig, 2004, S. 307) beschreibt drei wichtige Aspekte des Arbeitsgedächtnisses. Die erste Komponente, die *zentrale Exekutive* (central executive) ist ein System, welches den gesamten Ablauf kontrolliert, der zentralen Executive untergeordnet sind die *phonologische Schleife* und der *visuell-räumliche Notizblock*. Die wichtigste Aufgabe der *zentralen Executive* ist die Kontrolle der Aufmerksamkeit. So ermöglicht sie, dass Informationen ans Bewusstsein gelangen. Des Weiteren organisiert und kontrolliert sie die zu verarbeitende Informationen aus dem *visuell-räumlichen Notizblock* und der *phonologischen Schleife*. Die *phonologische Schleife* (phonological loop) dient der Speicherung und Kontrolle von sprachlichen Informationen, ist bedeutende Instanz beim Erlernen und Merken von Wörtern und Zahlen und weist große Ähnlichkeit mit dem Kurzzeitgedächtnis auf. Zudem besteht sie aus einem artikulatorischen Prozess, dieser ermöglicht die Aufrechterhaltung von Information durch Wiederholung. Der *visuell-räumliche Notizblock* (visuospatial scratch-pad) arbeitet auf ähnliche Weise wie auch die phonologische Schleife, ist jedoch vor allem für visuelle und räumliche Information verantwortlich (Baddeley, 1986, zitiert nach Zimbardo & Gerrig, 2004, S. 307).

2.2.6 LANGZEITGEDÄCHTNIS

Anders als im Kurzzeitgedächtnis werden Informationen im Langzeitgedächtnis unbegrenzt und dauerhaft gespeichert. Im Langzeitgedächtnis befindet sich unser gesamtes Wissen über die Welt und über uns selbst (Seel, 2003).

Der größte Teil der Inhalte, die in unserem Gedächtnis gespeichert sind, können dem Langzeitgedächtnis zugeordnet werden (Spada, 2006). Oftmals werden im Langzeitgedächtnis zwei Systeme, nämlich implizites bzw. prozedurales und explizites bzw. deklaratives Gedächtnis unterschieden (Zimbardo & Gerrig, 2004).

Tulving (1972) unterscheidet innerhalb des Langzeitgedächtnisses, das semantische und das episodische Gedächtnis, welche als deklaratives Gedächtnis zusammengefasst werden können. Deklarative Gedächtnisinhalte umfassen vor allem Faktenwissen, sind der Person bewusst und können verbalisiert werden (Tulving, 1972). Das semantische Gedächtnis umfasst all jene Bereiche, die mit Wissen einhergehen, beispielsweise sprachbezogenes Wissen oder Faktenwissen, das episodische Gedächtnis hingegen umfasst kontextbezogene Ereignisse des eigenen Lebens und ist erfahrungsbezogen (Tulving, 1972). Auch wenn die Differenzierung der beiden Systeme relativ gut gelingt, wird schnell klar, dass eine solche Unterscheidung praktisch schwieriger ist und das semantische und episodische Gedächtnis eng miteinander in Verbindung stehen und sich beeinflussen (Schermer, 2006).

Anders als im semantischen Gedächtnis sind die Inhalte des prozeduralen Gedächtnisses meist unbewusst und laufen automatisiert ab (Zimbardo & Gerrig, 2004), und ebenso fällt es schwer prozedurale Gedächtnisinhalte zu verbalisieren (Lefrancois, 2006).

Wie bereits zuvor erwähnt können Gedächtnisinhalte bzw. Wissen explizit bzw. implizit sein. Implizites oder auch nicht deklaratives Wissen ist meist unbewusst und kann schwer verbalisiert werden, explizite Gedächtnisinhalte können durch Anstrengung ins Bewusstsein gelangen (Lefrancois, 2006).

2.2.7 PHÄNOMENE WELCHE DEN ERWERB VON LERNINHALTEN BEEINFLUSSEN

Im folgenden Abschnitt wird auf jene Aspekte eingegangen, die einen Einfluss auf die Aneignung neuer Lerninhalte haben. Es werden einige Phänomene diskutiert, welche eine wesentliche Rolle in Bezug auf den Abruf von Lerninhalten spielen und diese beeinflussen können.

2.2.7.1 Positionseffekte: Primacy- und Recency- Effekt

Der serielle Positionseffekt kann ganz allgemein als eine Suche im Gedächtnis bezeichnet werden. Lerninhalte, welche am Anfang bzw. am Ende einer Lerneinheit eingeprägt werden, können besser erinnert und reproduziert werden als Lerninhalte im mittleren Bereich (Zimbardo & Gerrig, 2004).

Beim seriellen Positionseffekt kann einerseits der Primacy-Effekt und andererseits der Recency-Effekt unterschieden werden. Mit Primacy-Effekt ist die verbesserte Abruf- und Reproduktionsleistung von Lerninhalten am Anfang einer Lerneinheit gemeint. Im Gegensatz dazu wird mit dem Recency-Effekt die verbesserte Abruf- bzw. Reproduktionsleistung von Lerninhalten am Ende einer Lerneinheit verstanden (Zimbardo & Gerrig, 2004).

In einer frühen Studie von Jahnke (1965, zitiert nach Zimbardo & Gerrig, 2004, S. 312) sollten sich 48 StudentInnen Listen mit entweder 6, 10 oder 15 englischen Wörtern einprägen und sich an diese im Nachhinein erinnern. Die Hälfte der TeilnehmerInnen wurde aufgefordert, die Listen mit Wörtern sofort in derselben Reihenfolge in der sie auch präsentiert wurde, wiederzugeben. Der anderen Hälfte wurde mitgeteilt, es sollten so viele Wörter reproduziert werden, wie es für sie aus der Erinnerung möglich wäre. Die Ergebnisse zeigten, dass die Wiedergabeleistung für die Bedingung, in denen die StudentInnen so viele Wörter wie nur möglich erinnern sollten, besser war, als für die Bedingung, in der die Wörter in der Reihenfolge, in der sie auch präsentiert wurden erinnert werden sollten. Der Recency-Effekt konnte im Vergleich zum Primacy-Effekt stärker in der Bedingung der freien Reproduktion (es sollten so viele Wörter wie nur möglich erinnert werden) als für den Abruf der Wörter in derselben Reihenfolge (serial recall) beobachtet werden und außerdem stärker für kürzere Wortlisten als für längere (Jahnke, 1965, zitiert nach Zimbardo & Gerrig, 2004, S. 312).

Eine Erklärung für das Phänomen der seriellen Positionseffekte ist, dass am Anfang einer Lerneinheit die zu Beginn eingeprägten Lernelemente einer Wortliste in den noch gänzlich freien Kurzzeitspeicher gelangen und demnach auch stetig wiederholt werden können und somit die Übertragung der Lerninhalte in den Langzeitspeicher effizienter und wahrscheinlicher gelingt als es für Inhalte im mittleren Bereich einer Lerneinheit der Fall ist. Die Lerninhalte am Ende einer Lerneinheit können zwar nicht so oft wiederholt werden wie jene zu Beginn, jedoch sind sie zum Zeitpunkt der Wiedergabe noch im Kurzzeitspeicher vorhanden und der Zugriff und Abruf gelingt somit direkt durch das Kurzzeitgedächtnis (Schermer, 2006).

2.2.7.2 Kodierungsspezifität und Hinweisreize beim Abruf von Information

Der Kontext, in dem Informationen gelernt und eingeprägt werden, hat einen wesentlichen Einfluss darauf, wie gut das Wiedergeben bzw. der Abruf der gelernten Information gelingt (Schermer, 2006). Unter Enkodierungsspezifität versteht man, dass die Leistungen beim Wiedergeben von Informationen gesteigert werden können, wenn die Hinweisreize bzw. die Kontextbedingungen beim Einprägen der Informationen und beim Wiedergeben der Informationen die gleichen sind (Zimbardo & Gerrig, 2004).

Das Modell der Enkodierungsspezifität hat seinen Ursprung im Paarassoziationslernen. In diesem Zusammenhang waren es Wortpaare, in welchem das eine Wort als sogenannter Hinweisreiz fungiert, der die Wiedergabe des anderen Wortes erleichtert. Beim Paarassoziationslernen besteht das Material, welches gelernt werden soll zumeist aus zwei Lernelementen, beispielsweise rot- Hund. In der Wiedergabe- Phase wird das erste Lernelement vorgegeben (rot) und das zweite Lernelement aus der jeweiligen Kombination (Hund) soll wiedergegeben werden (Schermer, 2006). Ebenso wie beim Paarassoziationslernen, wo beispielsweise ein Wort in einer Wortkombination als Hinweisreiz vorhanden ist und die

Wiedergabeleistung erleichtert, können auch Gefühle und Stimmungen als Hinweisreize bzw. Variablen des Kontexts fungieren und die Wiedergabeleistung erhöhen, wenn der Inhalt in derselben Stimmung abgerufen wird, in der er auch eingeprägt wurde (Schermer, 2006).

In einer Studie von Balch und Lewis (1996) wurde der Einfluss des Kontextes auf die Gedächtnisleistung (Lernen von Wortlisten) an 168 StudentInnen überprüft. Den StudentInnen wurden zum Zeitpunkt des Einprägens der Wortlisten zwei Musikstücke (Mozart vs. Jazz) in jeweils einer Version mit langsamem Tempo und einer Version mit schnellem Tempo vorgespielt. Das Abrufen der Wortlisten erfolgte entweder im selben Kontext (gleiches Musikstück, gleiches Tempo), mit unterschiedlichem Tempo oder anderem Musikstück. Die Ergebnisse zeigten, dass die Gedächtnis- bzw. Reproduktionsleistung in der Bedingung, in der das Tempo der Hintergrundmusik im Vergleich zur Einprägephase gleich blieb am höchsten war, gefolgt von den Bedingungen, in denen sich das Tempo bzw. das Musikstück änderte (Bach & Lewis, 1996).

2.2.8 GESCHLECHTSUNTERSCHIEDE IM ZUSAMMENHANG MIT LERN- UND MERKFÄHIGKEIT

Die Frage nach kognitiven Leistungsunterschieden zwischen Männern und Frauen stand sehr oft im Mittelpunkt zahlreicher Untersuchungen (Halpern & LaMay, 2000). Im Folgenden seien einige Studien beschrieben, welche sich mit geschlechtsspezifischen Unterschieden in Bezug auf Gedächtnisleistungen beziehen.

In einer Studie von Colley, Ball, Kirby, Harvey und Vingelen (2002) wurde untersucht, ob Geschlechtsunterschiede in Bezug auf Gedächtnisleistungen beobachtet werden können. 60 weibliche und männliche Studenten wurden aufgefordert, Einkaufslisten auswendig zu lernen, die einmal als „Lebensmittel- Einkaufsliste“ und einmal als „Haushaltswaren-Einkaufsliste“ vorgegeben wurden. Desweiteren gab es drei unterschiedliche Instruktionsbedingungen. In der neutralen Bedingung wurde

den VersuchsteilnehmerInnen mitgeteilt, es handle sich bei dieser Untersuchung um eine Überprüfung der Gedächtnisleistung im Alltag, die alle Personen sehr häufig nutzen. In einer zweiten Instruktion wurden die TeilnehmerInnen daraufhin gewiesen, dass dies eine Untersuchung sei, welche die Gedächtnisleistung im Alltag überprüfe und bei der Frauen sehr gute Leistungen erzielen würden, und schließlich wurden die TeilnehmerInnen in einer dritten Bedingung instruiert, es ginge in dieser Studie um die Überprüfung der Gedächtnisleistung im Alltag und es sei bekannt, dass Männer in diesem Bereich gute Leistungen erzielen. Die Aufgaben umfassten in beiden Listen 16 Items, welche die Personen einmal sofort nach dem Einprägen reproduzieren sollten, danach bekamen sie eine Distraktionsaufgabe und im Anschluss daran sollten sie noch einmal versuchen, so viele Wörter wie möglich wiederzugeben. Es zeigte sich, dass beim direkten Wiedergeben mehr Items reproduziert werden konnten, als wenn eine Distraktionsaufgabe vorgegeben wurde. Wenn die Einkaufsliste unter dem Namen „Haushaltswaren-Liste“ vorgegeben wurde, zeigte sich kein Unterschied in der Wiedergabeleistung bei den Geschlechtern. Die Vorgabe der Einkaufsliste als „Lebensmittel-Liste“ ergab insgesamt für die weiblichen VersuchsteilnehmerInnen bessere Leistungen in der Reproduktion der Listen. Etwas schlechtere Reproduktionsleistungen der Studentinnen konnten in der Bedingung ermittelt werden, in der ihnen mitgeteilt wurde, dass Frauen über gute Gedächtnisleistungen verfügen und am niedrigsten war die Reproduktionsleistung der Frauen, wenn zugunsten der Männer instruiert wurde. Die Autoren postulieren an dieser Stelle eine mögliche Negativaktivierung von Stereotype Threat, im Gegensatz zu den männlichen Studenten, bei denen über alle drei unterschiedlichen Instruktionsbedingungen keine Unterschiede in der Gedächtnisleistung festgestellt wurden (Colley et al., 2002).

In einer Studie von Kramer, Delis, Kaplan, O'Donnel und Prifitera (1997) wurden Geschlechtsunterschiede in Bezug auf Gedächtnisleistung und verbalem Lernen über verschiedene Altersstufen hinweg erforscht. Den

Mädchen und Burschen im Alter zwischen fünf und sechzehn Jahren wurde ein an das Alter angepasster Vokabellertest und ein weiterer Verbaler Lerntest vorgegeben. Die Vorgabe des verbalen Lerntest erfolgte in fünf Durchgängen, nach jedem Durchgang wurden die VersuchsteilnehmerInnen aufgefordert, soviel wie möglich der zuvor präsentierten Wörter wiederzugeben. Im Anschluss wurde in einem Durchgang eine Interferenzliste vorgegeben. Danach wurden die Personen wiederum gebeten, die Listen, die ihnen mittels verbalem Lerntest vorgegeben wurden, einmal frei und einmal Mithilfe eines Hinweisreizes zu reproduzieren. In den darauffolgenden 20 Minuten bearbeiteten die weiblichen und männlichen TeilnehmerInnen Aufgaben, die nicht speziell verbales Lernen oder die Gedächtnisleitung betrafen. Im Anschluss daran wurden sie noch einmal aufgefordert, die zu Beginn gelernten Wortlisten des verbalen Lerntests einmal frei wiederzugeben und einmal mit Hilfe eines Hinweisreizes. Des Weiteren wurde nach zwanzig Minuten die Wiedererkennungs-Leistung der Wortlisten abgefragt, in der die Versuchspersonen jeweils zwei Wörter vorgegeben bekamen und entscheiden mussten, ob ihnen die Wörter bereits zu Beginn im Rahmen der Wortlisten im verbalen Lerntest vorgegeben wurden. Es zeigte sich, dass Mädchen in allen fünf Durchgängen des Lerntest sowie auch in der freien Reproduktion nach 20 Minuten bessere Leistungen erzielten als die männlichen Versuchsteilnehmer. In der Bedingung des Wiedererkennens nach 20 Minuten war zu beobachten, dass Mädchen bessere Diskriminationsleistungen aufwiesen, unter anderem deswegen, weil sie weniger oft fälschlicherweise angaben, Wörter aus der Ursprungsliste des Lerntests wiederzuerkennen (*false-positive errors*) als es die Burschen taten. Geschlechtsunterschiede in Bezug auf verbales Lernen und Gedächtnis konnte über alle untersuchten Altersstufen hinweg beobachtet werden, tendenziell jedoch nehmen diese Geschlechtsunterschiede mit zunehmendem Alter zu (Kramer et al., 1997).

Zu ähnlichen Ergebnissen kamen auch West, Crook und Barron (1992), die neben dem Einfluss von Alter und individuellen Differenzen auch

Geschlechtsunterschiede in Bezug auf die Gedächtnisleistung überprüfen. Einer sehr großen Stichprobe von 2495 Versuchspersonen im Alter zwischen 18 und 90 Jahren wurde eine umfangreiche Testbatterie zur Erfassung der Gedächtnisleistung vorgelegt. Diese beinhaltete unter anderem sofortige und zeitverzögerte Wiedergabe 30 Minuten nach Vorgabe des Lerntest, Wiedergabe mit Hilfe von Hinweisreizen und eine Bedingung, in der die Leistung beim Wiedererkennen erfasst wurde. Es zeigten sich Leistungsunterschiede zugunsten der Frauen in allen Reproduktionsbedingungen im Untertest, in dem man sich eine Liste von Vor- und Nachnamen merken musste und im Untertest, der das Auswendiglernen einer Lebensmittel-Liste erforderte. Die Autoren gehen jedoch davon aus, dass die resultierenden Leistungsunterschiede zugunsten der Frauen einerseits durch die Erfahrung bzw. Übung im Umgang mit Lebensmittel-Listen bzw. aufgrund der Motivation der Frauen ein gesellschaftliches Stereotype zu bestätigen, nämlich beispielsweise als sozial kompetenter zu gelten als Männer und dadurch im Untertest Namen-Lernen besser abschneiden, relativ gesehen werden müssen (West et al., 1992).

In einer Studie von Ionescu (2000) wurden Geschlechtsunterschiede in Bezug auf visuelle und verbale Merkfähigkeit an PsychologiestudentInnen analysiert. Zusätzlich wurde geprüft, inwiefern sich die Geschlechter in der Einschätzung ihrer Leistung unterscheiden. Aus diesem Grund sollten die VersuchsteilnehmerInnen einmal vor dem Lernen und einmal nach der Reproduktion der 80 visuellen und verbalen Stimuli ihre eigene Leistung einschätzen. Die Ergebnisse der Untersuchung ergaben keine Unterschiede in der Reproduktionsleistung zwischen Männern und Frauen, weder für visuelles (Bilder) noch für verbales (Wörter) Testmaterial. Die Einschätzungen vor der Bearbeitung der Merkfähigkeitsaufgaben entsprachen sowohl bei Männern als auch bei Frauen ziemlich genau den Ergebnissen, die ermittelt werden konnten. Jedoch schätzten sich beide Geschlechter nach der Reproduktion schlechter ein als die resultierenden Ergebnisse, Männer schätzten sich

schlechter ein, was visuelles Testmaterial betraf, Frauen hingegen unterschätzten ihre Leistung in Bezug auf verbales Testmaterial.

In einer weiteren Studie von Ionescu (2002) wurden Geschlechterunterschiede in Bezug auf visuelle als auch verbale Merkfähigkeitsleistung und deren Einschätzung anhand einer größeren Stichprobe ($n = 207$) untersucht. Wiederum wurden den VersuchsteilnehmerInnen, die allesamt StudentInnen der Psychologie waren, 80 Stimuli (40 Wörter und 40 Bilder) zum Einprägen vorgegeben und wurden im Anschluss daran aufgefordert so viele Wörter bzw. Bilder wie möglich zu reproduzieren. Zusätzlich wurden die VersuchsteilnehmerInnen instruiert, die eigene Leistung vor und nach der Bearbeitung einzuschätzen, indem sie die Zahl der Wörter und der Bilder angeben, von der sie glaubten diese auch wiedergeben zu können. Auch hier konnten keine Geschlechtsunterschiede in Bezug auf die Reproduktionsleistung von visuellem und verbalem Testmaterial zwischen Männern und Frauen gefunden werden. Hinsichtlich der Einschätzung der Reproduktionsleistung konnte beobachtet werden, dass sich beide Geschlechter sowohl vor als auch nach der Bearbeitung des Testmaterials in Bezug auf ihre Leistung unterschätzten, Frauen schätzten sich jedoch im Vergleich zu den Männern signifikant öfter schlechter ein als die tatsächlichen Ergebnisse, die ermittelt werden konnten (Ionescu, 2002).

Zu ähnlichen Ergebnissen kamen Trahan und Quintana (1990), die in ihrer Untersuchung anhand von 140 VersuchsteilnehmerInnen Geschlechtsunterschiede in Bezug auf visuelle und verbale Gedächtnisleistung zu ermitteln versuchten. Auf Grundlage von vier Verfahren wurden visuelle und verbale Gedächtnisleistungen in Form von Wiedererkennung, sofortige Wiedergabeleistung, als auch verzögerte Abrufleistung erfasst. Es zeigten sich keine Unterschiede in Bezug auf visuelle oder verbale Gedächtnisleistung zwischen Männern und Frauen, einzig das Lernen von Wortlisten konnte zugunsten der Frauen ermittelt werden (Trahan & Quintana, 1990).

In einer Studie von Herrmann, Crawford und Holdsworth (1992) wurde der Einfluss der Vorgabe von Geschlechtsstereotypen Testmaterial auf die Gedächtnisleistung von Männern und Frauen untersucht. Einmal wurden die VersuchsteilnehmerInnen aufgefordert, ein für das weibliche Stereotyp entsprechendes Testmaterial zu lernen, nämlich sogenannte Einkaufslisten und einmal wurden die TeilnehmerInnen aufgefordert, ein für das männliche Stereotyp entsprechendes Testmaterial zu lernen, nämlich das Lernen von Wegbeschreibungen. Übereinstimmend mit dem jeweiligen Geschlechtsstereotype zeigten Frauen bessere Reproduktionsleistungen, wenn es sich um das Lernen von Einkaufslisten handelt, Männer hingegen zeigten bessere Merkfähigkeitsleistungen, sobald es um das Reproduzieren von Wegbeschreibungen geht (Herrmann et al., 1992).

In einer Studie von Choi und L'Hirondelle (2005) wurden Geschlechterunterschiede in Bezug auf die Lokalisation von Objekten untersucht (object location memory). Männlichen und weiblichen Testteilnehmern wurde ein Test zur Lokalisation von Objekten vorgegeben und zufällig einer der beiden Bedingungen mit abstrakten oder konkreten Objekten zugeteilt. Entsprechend den Vermutungen der Autoren konnte ermittelt werden, dass die Richtung der Geschlechtsunterschiede kontextspezifisch ist, demnach von der Art der Vorgabe des Testmaterials abhängt und ein Vorteil zugunsten der Frauen nur in der Bedingung, in welcher konkrete Objekte vorgegeben wurden, festgestellt werden konnte. Als ein Grund für den Vorteil von Frauen in Bezug auf die Lokalisation von konkreten Objekten, kann eine Überlegenheit der Frauen in verbalen Gedächtnisleistungen angenommen werden (Choi & L'Hirondelle, 2005).

2.3 STEREOTYPE UND STEREOTYPE THREAT

In diesem Kapitel wird das Phänomen Stereotype Threat anhand aktueller Forschungsergebnisse beschrieben, insbesondere die Auswirkungen von Stereotype Threat auf die Merkfähigkeit und hier vor allem im Zusammenhang mit dem Geschlecht.

Des Weiteren sollen Studien beschrieben werden, welche sich mit den Möglichkeiten der Reduzierung von Stereotype Threat beschäftigen.

2.3.1 DEFINITION VON STEREOTYPE

Die Zuschreibung von Eigenschaften gegenüber einer Menschengruppe, nennt man Stereotype und diese dienen im Grunde dazu, unsere Sicht von einer komplexen Welt zu vereinfachen und überschaubar zu machen (Aronson, Wilson & Akert 2008; Herkner, 2003).

Petersen & Six (2008) gehen davon aus, dass für die Entstehung von Stereotypen bei Menschen eine generelle Affinität zur Bildung von sozialen Kategorien besteht. Diese Kategorisierung von Menschen in Gruppen, die für die Entstehung von Stereotypen essentiell ist, kann von weit umfassenden Kategorien von Merkmalen (z.B. Geschlecht, Nationalität) bis hin zu kleineren Kategorien (z.B. Karrierefrauen) reichen.

Klauer (2008) geht davon aus, dass soziale Kategorien inhaltlich mit spezifischen Erwartungen, die man hinsichtlich der Eigenschaften und Verhaltensweisen an die Mitglieder einer Gruppe hat, verbunden sind. Aus diesen Erwartungen entstehen sozial geteilte Wissensstrukturen, die als kognitive Schemata oder Stereotype bezeichnet werden und simplifizierenden Verarbeitungs- und Urteilsheuristiken sehr ähnlich sind.

Sind Stereotype einmal vorhanden, nehmen sie Einfluss auf die Prozesse der Aufmerksamkeit, die Interpretation von Informationen, auf das Gedächtnis und das schlussfolgernde Denken und beeinflussen somit Informationsverarbeitungsprozesse (Petersen & Six, 2008).

Stereotype haben nicht nur eine Wirkung auf denjenigen, der sie wahrnimmt, sondern beeinflussen auch jene Personen, die Mitglied der stereotypisierten Gruppe sind. Zum einen können sich diese Personen als Reaktion auf das Verhalten des Mitgliedes einer Fremdgruppe dem Stereotyp entsprechend verhalten und somit das ursprünglich angenommene Stereotyp bestätigen oder es kann alleine dadurch, dass den Mitglieder einer stereotypisierten Gruppe bewusst ist, dieser Gruppe

anzugehören, das stereotyp konforme Verhalten ausgelöst werden (Petersen & Six, 2008).

2.3.2 BEDROHUNG DURCH STEREOTYPE- STEREOTYPE THREAT

Im Folgenden wird ein spezieller Bereich im Rahmen von Stereotypen beschrieben, nämlich das Phänomen *Stereotype Threat*.

Es gibt zahlreiche Untersuchungen, die sich mit vorurteilsbehafteten Personen, Stereotypen und Diskriminierung von Gruppen beschäftigt haben. Im Gegensatz dazu wurde in wissenschaftlichen Untersuchungen relativ selten der Frage nachgegangen, wie denn Stereotype auf diejenigen Personen wirken, die direkt davon betroffen sind (Bierhoff, 2000). Ein besonders bedeutsamer Bereich, bei dem negative Stereotype eine Wirkung auf die Personen der stereotypisierten Gruppe zeigen können, ist der akademische Leistungsbereich, wie beispielsweise universitäre Aufnahmeverfahren oder Tests im Rahmen der Personalauswahl (Keller, 2008).

Es existiert eine Reihe von Erklärungen, warum beispielsweise asiatische Amerikaner in Leistungs- und Intelligenztests etwas besser abschneiden als Angloamerikaner und warum wiederum Afroamerikaner etwas schlechter als Angloamerikaner (Aronson et al., 2008; Bierhoff, 2000). In der Regel wurde aber der Rolle der Testsituation als einer der Ursachen, die sich ungünstig auf die Leistung auswirken kann, nur geringe Aufmerksamkeit geschenkt (Bierhoff, 2000).

In der Stereotype-Threat-Theorie (Steele & Aronson, 1995) geht man davon aus, dass Personen sich bedroht fühlen, sobald sie sich in einer Situation wiederfinden, in der sie befürchten müssen, einer Beurteilung ausgesetzt zu werden, die aufgrund von negativen Stereotypen zustande gekommen ist oder das Stereotyp der Gruppe, der sie angehören, durch ihr eigenes Verhalten unabsichtlich zu bestätigen. Das Phänomen, in Situationen eine Bedrohung durch negative Stereotype zu erleben, wird „Stereotype Threat“ genannt (Steele & Aronson, 1995).

Man geht demnach davon aus, dass sich negative Stereotypen auf die Testleistung einer Person auswirken können, und zwar insofern, als dass die Person in diesem Fall geringere Leistungen zeigt, als es ihr eigentliches Leistungsvermögen zulassen würde (Keller, 2008).

In einer der zahlreichen Untersuchungen von Steele & Aronson (1995) wurde das Phänomen der Bedrohung durch Stereotype analysiert. Mit den Probanden, die allesamt Studierende an einer Universität und entweder Afroamerikaner bzw. Weiße waren, führten die Untersucher jeweils einzeln einen mündlichen Test durch. Jeweils der Hälfte der weißen und afroamerikanischen StudentInnen wurde mitgeteilt, das Ziel der Studie sei, intellektuelle Fähigkeiten zu erfassen. Der anderen Hälfte der StudentInnen wurde lediglich gesagt, dass die Forscher versuchen, einen verlässlichen Test zu entwickeln und dies der Grund für die Testdurchführung sei, man könne anhand der Testergebnisse aber keine Rückschlüsse über die tatsächlichen intellektuellen Fähigkeiten der TeilnehmerInnen machen. Die Ergebnisse zeigten, dass weiße Studierende, gleichgültig, ob man ihnen sagte, bei der Testdurchführung gehe es nur darum einen Test zu entwickeln bzw. man ihnen mitteilte, der Test solle ihre intellektuellen Fähigkeiten erfassen, jeweils in beiden Bedingungen gleich gute Leistungen zeigten. Die Gruppe der afroamerikanischen Studierenden, denen man mitgeteilt hatte, die Ergebnisse des Tests würden nur dazu verwendet werden, ein geeignetes Verfahren zu entwickeln, schnitten genauso gut ab wie auch die weißen StudentInnen. Nur jene Gruppe der Afroamerikaner, die annahm, ihre Fähigkeiten wurden gemessen, schnitt deutlich schlechter ab als die weißen und afroamerikanischen Probanden der anderen Bedingungen (Steele & Aronson, 1995).

2.3.2.1 Stereotype Threat und Geschlecht

Wie sich in einer Studie von Spencer, Steele & Quinn (1999) zeigte, kann das Phänomen der Bedrohung durch Stereotype auch auf das Geschlecht übertragen werden. In dieser Untersuchung wurden teilnehmende Frauen

einmal dahingehend instruiert, es ginge bei der Studie darum, Unterschiede der mathematischen Fähigkeiten zwischen Männern und Frauen herauszufinden, beim anderen Mal wurde ihnen mitgeteilt, der mathematische Test habe nichts mit Geschlechterunterschieden zu tun. Auch hier zeigte sich, dass Frauen dann gleich gut wie ihre männlichen Kollegen abschnitten, wenn man ihnen sagte, der Test habe nichts mit Geschlechterunterschieden zu tun. Schlechtere Ergebnisse als die Männer erzielten sie, sobald man sie auf Geschlechterunterschiede bei mathematischen Fähigkeiten aufmerksam machte (Spencer et al., 1999).

Um die Effekte von Stereotype Threat auf die schulische Leistung evaluieren, gab man in einer Studie von Keller und Dauenheimer (2003) Schülern und Schülerinnen der 10. Schulstufe einen Mathematiktest vor. Durch randomisierte Zuweisung erhielten die SchülerInnen eine der beiden Testversionen von Mathematikaufgaben. In der ersten Version erfuhren die Probanden vor der Bearbeitung, dass es sich hierbei um Aufgaben handle, bei denen man in vorherigen Untersuchungen bereits Geschlechtsunterschiede hinsichtlich der Leistung feststellen hatte können. Bei der zweiten Version des Tests gab man den SchülerInnen die Information, dass bei dieser Art der Aufgaben bislang noch keine Geschlechtsunterschiede festzustellen waren. Mit Ausnahme dieser zwei unterschiedlichen Informationen zu Beginn des Tests waren die Aufgaben der beiden Bedingungen völlig gleich. Die Ergebnisse machen deutlich, dass Mädchen sehr viel bessere Leistungen zeigten, wenn man vorab die Information gab, dass sich beim Lösen der Aufgaben keine Unterschiede zwischen den Geschlechtern gezeigt hatten. Die Schülerinnen waren nun der Überzeugung, dass ihr Geschlecht ihre Leistung in keiner Weise beeinflussen würde und erzielten gleich gute Leistungen wie ihre männlichen Kollegen. Im Gegensatz dazu kam das Geschlechtsstereotyp in der zweiten Bedingung zum Tragen und dies spiegelte sich auch in den Ergebnissen wider. Mädchen zeigten schlechtere Leistungen als Buben, wenn Geschlechtsstereotype in der Situation hervorstechend und anwendbar war. Die Ergebnisse dieser Untersuchung weisen darauf hin,

dass sich nur eine geringe Variation der Testbedingungen stark auf die Leistungsergebnisse bei Mädchen auswirken kann (Keller & Dauenheimer, 2003).

2.3.2.2 Stereotype Threat und Ängstlichkeit

In einer Untersuchung von Marx und Stapel (2006) wurde Frauen und Männern ein Test mit mathematischen Aufgaben vorgelegt, worin jene Frauen, bei denen zuvor ein negatives Stereotype aktiviert worden war, deutlich schlechter abschnitten als die restlichen Probanden. Zusätzlich wurde von den Untersuchern vor Beginn des Tests die Ängstlichkeit und im Anschluss an den Test die Frustration erhoben. Auch hier fielen die Ergebnisse wie erwartet aus. Jene Frauen, bei denen vor Beginn des Tests ein Geschlechterstereotyp aktiviert worden war, zeigten bedeutend höhere Werte in der Skala Ängstlichkeit vor Beginn des Tests sowie höhere Frustrationswerte nach dem Lösen der Testaufgaben im Vergleich zu jener Gruppe der Frauen, bei der keine negative Stereotype aktiviert worden war (Marx & Stapel, 2006).

Ähnliche Ergebnisse konnte man auch in einer Studie von Stone, Lynch, Sjomeling & Darley (1999) feststellen. Bei den teilnehmenden Personen wurde vor und nach der Bearbeitung eines Tests die Ängstlichkeit erfasst. Die Ergebnisse belegten, dass jene Personen, die in der Stereotype Threat Bedingung waren, nach dem Test höhere Angstwerte aufwiesen als jene in der Bedingung ohne Stereotype Threat Aktivierung (Stone et al., 1999).

2.3.2.3 Stereotype Threat und Entscheidungsverhalten

In einer Untersuchung von Carr und Steele (2009) wurde der Einfluss von Stereotype Threat auf die Problemlösefähigkeit untersucht. Es wurde angenommen, dass Personen unter Stereotype Aktivierung das Aufgeben alter, zuvor erfolgreicher Lösungsstrategien und in weiterer Folge die Anwendung neuer, effizienterer Strategien, wenn sich die Bedingungen ändern, schwerer fallen als Personen ohne Stereotype Threat Aktivierung.

In der Stereotype Threat Bedingung wurde den TeilnehmerInnen mitgeteilt, einen Test zu bearbeiten, welcher höchst valide mathematische Fähigkeiten überprüfe. In der Kontrollgruppe wurde ihnen gesagt, es handle sich um ein Verfahren, welches herausfinden solle, wie genau Puzzles von Personen gelöst werden. Tatsächlich wurde allen Personen ein Verfahren zur Bearbeitung vorgegeben, um die Ausdauer in Bezug auf das Problemlösen zu erfassen. Die zentrale Hypothese war, dass durch die Auslösung von Stereotype Threat eine unflexible Beharrlichkeit im Denken begünstigt wird, somit alte Strategien nicht aufgegeben und neue effizientere Strategien nicht gewählt werden, auch wenn die Situation es erfordert. Zusätzlich wurde nach der Testvorgabe zur Stereotype Unterdrückung ein lexikaler Entscheidungstest vorgegeben, in dem die TeilnehmerInnen auf einem Computerbildschirm eine Liste von Buchstaben dargeboten bekamen und sich entscheiden sollten, ob ein Wort schon gezeigt worden war oder noch nicht. Eine Liste von Wörtern wurde dargeboten, die mit negativen Stereotypen in Bezug auf Frauen und Mathematik verbunden waren. Die Ergebnisse zeigen, dass Stereotype Threat die Wahrscheinlichkeit von unflexibler Ausdauer in Bezug auf Problemlösestrategien begünstigt (Carr & Steele, 2009).

2.3.2.4 Stereotype Threat und Gedächtnis

In einer Untersuchung von Schmader und Johns (2003) wurde der Einfluss einer negativen Aktivierung durch Stereotype Threat auf die Kapazität des Arbeitsgedächtnisses untersucht. An ihrer ersten Studie nahmen 75 männliche und weibliche Psychologiestudenten teil. Überprüft wurde, ob die Leistungen der Arbeitsgedächtnis-Kapazität der weiblichen Teilnehmerinnen unter der Stereotype Threat Bedingung abnahmen und sich von den Leistungen der männlichen Studenten und der weiblichen Kolleginnen ohne Stereotype Threat Instruktion unterschieden. Unter dem Vorwand der Vorgabe eines Tests zur Erfassung der mathematischen Fähigkeiten wurde die Arbeitsgedächtnis- Kapazität untersucht, indem die Versuchsteilnehmer nach der Vorgabe einer mathematischen Gleichung

ein Wort vorgegeben bekamen. Im Anschluss an die Bearbeitung der Aufgaben wurden die TeilnehmerInnen aufgefordert, so viele Wörter wie möglich zu reproduzieren, an die sie sich erinnern konnten. Wie erwartet zeigten die Ergebnisse, dass sich bei Frauen, die den Test unter Stereotype Threat Bedingung bearbeiteten, ein negativer Einfluss hinsichtlich der Gedächtnis-Kapazität, gemessen an der Zahl der Wörter, welche im Anschluss an den mathematischen Test wiedergegeben werden konnten, zeigte. Im Vergleich zu den männlichen Studenten und den weiblichen Studentinnen ohne Stereotype Threat Instruktion zeigten weibliche Teilnehmerinnen unter Stereotype Threat Bedingung schlechtere Gedächtnisleistungen.

Zu ähnlichen Ergebnissen kamen Schmader und Johns (2003) auch in ihrer zweiten Studie, in der sie weiße und lateinamerikanische StudentInnen in Bezug auf die Gedächtnisleistung unter Stereotype Threat Bedingungen untersuchten. Einer Gruppe der VersuchsteilnehmerInnen wurde mitgeteilt, dass es sich um ein Testverfahren handle, welches die allgemeine Intelligenz erfasse. Die Ergebnisse zeigten, dass auch in dieser Untersuchung die lateinamerikanischen StudentInnen unter Stereotype Threat Bedingung schlechtere Gedächtnisleistungen aufwiesen (gemessen an der Anzahl der Wörter, die wiedergegeben werden konnten) als ihre weißen MitstudentInnen und die lateinamerikanischen StudentInnen der Kontrollgruppe.

In der dritten und letzten Untersuchung von Schmader und Johns (2003) ging man im Rahmen eines standardisierten akademischen Tests der Frage nach, ob die Arbeitsgedächtniskapazität die Stereotype Threat Effekte vermittelt. Im Gegensatz zu den beiden vorangegangenen Untersuchungen erfolgte die Überprüfung der Arbeitsgedächtnisleistung mittels eines Arbeitsgedächtnistest, in dem die Personen die Vokale in einem Satz, den sie vorgegeben bekamen, zählen und im Anschluss an jeden Satz ein Wort im Satz reproduzieren sollten. Weiters gab es wieder eine Aktivierung von Stereotype Threat. Frauen in der Stereotype Threat Bedingung bearbeiteten den Test zur Überprüfung des Arbeits-

gedächtnisses in einem Raum mit zwei weiteren männlichen Kollegen und einem männlichen Versuchsleiter, die Frauen in der Kontrollbedingung hingegen bearbeiteten den Arbeitsgedächtnistest mit zwei Kolleginnen und einer weiblichen Versuchsleiterin. Zur Vorgabe des Mathematiktests befanden sich alle Versuchsteilnehmerinnen in einem Raum. Der Kontrollgruppe wurde mitgeteilt, es handle sich bei dem Verfahren um einen in Entwicklung befindlichen Problemlösetest und den Personen in der Stereotype Threat Bedingung wurde mitgeteilt, der Test sei ein Verfahren zur Überprüfung der mathematischen Begabung. Die Ergebnisse zeigten, dass Frauen in der Stereotype Threat Bedingung weniger Wörter im Gedächtnistest wiedergeben konnten als Frauen in der Kontrollbedingung. Die Ergebnisse der letzten Studie machen deutlich, dass die verminderte Gedächtnisleistung unter Stereotype Threat vermittelnd auf die ebenso verminderten Leistungen im standardisierten Mathematiktest wirkt (Schmader & Johns, 2003).

2.3.2.5 Reduzierung von Stereotype Threat

In einer Untersuchung von Rydell, Conell & Beilock (2009) wurde der Frage nachgegangen, ob eine alternative Denkweise bzw. ein positives Stereotype die Leistung verbessern kann und so die Bedrohung durch Stereotype verringert wird. Die weiblichen Studentinnen, welche an der Untersuchung teilnahmen, erhielten einen Mathematiktest und bekamen zufällig eine der vier Bedingungen des Tests. In der ersten Bedingung erhielten die Probanden keine Information. In einer weiteren Bedingung lasen die Teilnehmerinnen, dass es laut Literatur bekannt sei, dass Frauen schlechtere mathematische Leistungen zeigen als Männer. In einer dritten Version wurde den Probanden erklärt, warum Collegestudenten bessere mathematische Leistungen vorzuweisen haben als Nicht-Collegestudenten; hiermit sollte ein positives Stereotype aktiviert werden. In der letzten Bedingung wurden die Teilnehmer darüber informiert, warum weibliche Personen schlechter im Lösen von Mathematikaufgaben seien als Männer und warum Collegestudenten in

der Regel bessere Leistungen bei Mathematikaufgaben zeigen als Nicht-Collegestudenten. In der letzten Manipulation wurde den Studentinnen somit sowohl ein negatives Stereotyp als auch ein positives Stereotype präsentiert. Wie zu erwarten zeigte sich, dass die Leistungen der Studentinnen, die mit einem negativen Stereotyp konfrontiert wurden, schlechter ausfielen als bei jener Gruppe, der man keine Information über das Vorhandensein von Stereotypen gab. Ein weiteres interessantes Ergebnis dieser Untersuchung war, dass jene Studentinnen, die man über ein negatives (Frauen zeigen schlechtere Leistungen in Mathematik als Männer) und ein positives Stereotype (Collegestudenten zeigen bessere Leistungen im Lösen von Mathematikaufgaben als Nicht-Collegestudenten) informierte, bessere Leistungen zeigten als jene, bei denen nur eine Aktivierung des negativen Stereotypes stattfand. Sobald sich die Studentinnen für ein negatives oder positives Stereotyp ihrer sozialen Gruppe entscheiden konnten, wählten sie das positive und dies hatte wiederum einen positiven Effekt auf ihre Leistungen (Rydell et al., 2009).

Zu ähnlichen Ergebnissen kamen auch McGlone und Aronson (2006). Bevor sie die Teilnehmer einen Test zum räumlichen Vorstellungsvermögen bearbeiten ließen, erinnerten sie die Probanden daran, dass sie Studierende einer ausgewählten Privatuniversität seien. Diese Information reichte aus, um den Unterschied zwischen weiblichen und männlichen Studierenden verschwinden zu lassen, denen sie nur mitteilten, Bewohner des Nordostens zu sein. So konnte auch hier gezeigt werden, dass negative Stereotype (Frauen sind schlecht in mathematischen Aufgaben) durch positive Stereotype (Ich bin ein ausgewählter Student) ausgetauscht werden konnte und sich dies positiv auf die Leistungen der weiblichen Studentinnen auswirkte (McGlone & Aronson, 2006).

In einer Studie von Martens, Greenberg und Schimel (2006) wurden die Auswirkungen von Selbstbestätigung auf die Leistung von Frauen unter Stereotype Threat Bedingung an 77 Frauen und 70 Männern untersucht. In der ersten Studie konnte gezeigt werden, dass Frauen beim Bearbeiten eines Mathematiktests schlechter abschnitten, wenn ihnen mitgeteilt

wurde, der Test sei ein diagnostisches Instrument, um die mathematische Intelligenz zu erfassen (Stereotype Threat Bedingung). In einer dritten Bedingung wurde Frauen, welche der Stereotype Threat Bedingung zugeordnet wurden, zusätzlich eine Liste mit Eigenschaften vorgegeben, die sie in Bezug auf persönliche Wichtigkeit rangreihen sollten. Dabei waren Fragen bezüglich der Beziehung zu Freunden und Familie oder zur physischen Attraktivität zu beantworten. Im Anschluss daran wurden die Frauen unter Stereotype Threat/Selbstbestätigungs-Bedingung aufgefordert, jene Charaktereigenschaft anzugeben, welche sie am meisten an sich selbst schätzen würden und ihnen wichtig sei. Die Ergebnisse zeigten, dass Frauen in der dritten Bedingung (Stereotype Threat/Selbstbestätigungs-Bedingung) gleich gute Leistungen im Bearbeiten des Mathematiktests aufwiesen wie auch Frauen in der Bedingung ohne Stereotype Threat. Somit kann angenommen werden, dass durch die Selbstbestätigungs-Angaben vor Beginn des Tests, die mathematischen Leistungen der Frauen verbessert werden können und so Auswirkungen der Stereotype Threat – Bedingung verschwinden. In einer zweiten Studie nahmen 55 männliche und 55 weibliche Studentinnen teil, welche zufällig einmal einer Stereotype Threat Bedingung, in der ihnen mitgeteilt wurde, dass Frauen schlechter bei Tests zum räumlichen Vorstellen abschneiden und einmal einer Kontrollbedingung (ohne Stereotype Threat) zugeteilt wurden. Die TeilnehmerInnen wurden aufgefordert, einen Test zum räumlichen Vorstellungsvermögen zu bearbeiten. Etwa der Hälfte der Frauen und der Hälfte der Männer wurde in der Stereotype Threat Bedingung vor dem Bearbeiten des Tests eine Selbstbestätigungsfrage gestellt. Frauen zeigten bessere Leistungen im Test zum räumlichen Vorstellen, sofern sie sich in der Stereotype Threat/ Selbstbestätigungs-Bedingung befanden. Bei ihren männlichen Mit-Studenten war hingegen keine Verbesserung der Leistung festzustellen, wenn sie in der Selbstbestätigungsbedingung waren (Martens et al., 2006).

Eine andere Möglichkeit der Reduzierung der Auswirkung von Stereotype Threat, nämlich die Aufklärung über das Phänomen Stereotype Threat

untersuchten Johns, Schmader und Martens (2005). In ihrer Untersuchung wurde 144 StudentInnen der Statistik ein Mathematiktest vorgegeben, der einer Gruppe als Problemlösetest und der anderen Gruppe als Mathematiktest. In einer dritten Bedingung wurde den TestteilnehmerInnen ebenso mitgeteilt, es handle sich um einen Mathematiktest, nur wurden die TeilnehmerInnen dieser Gruppe zusätzlich über das Phänomen Stereotype Threat und dessen möglicherweise ungünstigen Einfluss auf die Leistungen von Frauen im Test aufgeklärt. Wie erwartet zeigte sich, dass Frauen, wenn der Test lediglich als Mathematiktest beschrieben wurde, ohne über das Phänomen Stereotype Threat aufzuklären, schlechtere Leistungen im Test zeigten als ihre männlichen Kollegen. Die Leistungen jener Frauen in der Bedingung, in der der Test als Problemlösetest instruiert wurde und in der Bedingung, in der die Frauen über das Phänomen Stereotype Threat aufgeklärt wurden, unterschieden sich nicht von den Leistungen der Männer. Des Weiteren konnte festgestellt werden, dass die Leistungen der Frauen in der Bedingung der Aufklärung des Phänomens umso eher stiegen, je mehr diese ihre Angst dem Phänomen Stereotype Threat zuschrieben. Im Gegensatz dazu verschlechterten sich die Leistungen der Frauen in der Mathematiktest Bedingung, je mehr diese ihre Angst dem Stereotyp Threat zuschrieben (Johns et al., 2005).

In einer Studie von Ambady, Paik, Steele, Owen-Smith and Mitchell (2003) wurden die Effekte der Stärkung der persönlichen Individualität (*Individuation*) auf das Phänomen von Stereotype Threat untersucht und der Frage nachgegangen, ob durch die *Individuation* vor der Bearbeitung eines Tests unter Stereotype Threat Bedingung die Aktivierung bzw. die Identifizierung mit dem instruierten Stereotype abgeschwächt werden konnte oder ganz verschwand und sich die Testleistungen verbesserten. An der Untersuchung nahmen 45 weibliche Studentinnen teil. Der Hälfte der Teilnehmerinnen wurde vor der Testvorgabe eine Liste von Wörtern vorgegeben, welche dem weiblichen Geschlecht entsprachen, wie beispielweise Kleid, Mutter oder Ohrring. Die andere Hälfte der

Teilnehmerinnen wurde neutral instruiert. In jeder dieser beiden Gruppen wurde wiederum jeweils der Hälfte der Teilnehmerinnen ein Individualitäts-Fragebogen vorgelegt, während bei dem anderen Teil der Fragebogen inhaltlich neutral war. Die Ergebnisse zeigen, dass jene Gruppe, die nicht geschlechtsspezifisch instruiert wurde und keine Angaben zur *Individuation* machten (Kontrollgruppe) besser abschnitten als jene Gruppe, die in Richtung Geschlecht instruiert wurde und keine Informationen zu persönlicher Individualität angeben mussten. Jene Gruppe, welche in Richtung Geschlecht instruiert wurde und Angaben zur Selbstwertung machte, zeigten bessere Testleistungen als jene Gruppe, die nur in Richtung Geschlecht instruiert wurden ohne Information zur Individualität zu machen (Ambady et al., 2003).

2.4 RISIKOVERHALTEN

2.4.1 DEFINITION VON RISIKOVERHALTEN

„Risikoverhalten kann in einer sehr allgemeinen Form als Verhalten in Wahlsituationen mit ungewissem Ausgang betrachtet werden“ (von Klebelsberg, 1969, S.44).

Diese Definition bezieht sich auf eine ganze Reihe relativ unterschiedlicher Bereiche, in denen man Risikoverhalten beobachten und untersuchen kann, wie beispielsweise Rauchen, Drogenkonsum, ungeschützter Geschlechtsverkehr, etc. (Lerch, Rauch & Schlesier, 2000).

Risiko bedeutet in diesem Zusammenhang ganz allgemein die Folge einer Handlung oder eines Ereignisses, die mit einem Schaden oder Verlust in Zusammenhang gebracht werden können (Jungerman & Slovic, 1993, zitiert nach Lerch et al., 2000, S.17). Der Begriff des Verlustes muss hier relativ gesehen werden, weil es bei der Entscheidung für eine Handlungsalternative immer um den subjektiv empfundenen Verlust, der eintreten könnte und um subjektives Abwägen der Für und Wider eines Nutzens der Person geht (Lerch et al., 2000). Um dies mit einem Beispiel

zu veranschaulichen, könnte man also annehmen, dass eine Person nur dann betrunken mit dem Auto fährt, wenn sie sich subjektiv einen Nutzen davon erwartet, eventuell auf fragliche Art und Weise einer Frau zu imponieren, denn ansonsten würde die Person das Risiko, sich selbst bzw. andere durch das Verhalten zu gefährden wohl kaum eingehen.

Ein weiterer Bereich der psychologischen Risikoforschung, neben dem des Gefahrenrisikos (Rauchen, ungeschützter Geschlechtsverkehr), stellt das kognitive Risiko dar. Das sogenannte kognitive Risiko ist im Bereich von schulischer Lernleistung und Leistungsprozessen von großem Interesse, vor allem hinsichtlich der Evaluation von Lernerfolgen durch schulische Leistungstests mit Mehrfachwahlantworten, das sogenannte Multiple Choice Antwortformat (Lerch et al., 2000).

2.4.1.1 Definition von Entscheidung

„Mit dem Begriff Entscheidung verbindet man im Allgemeinen mehr oder weniger überlegtes, konfliktbewusstes, abwägendes und zielorientiertes Handeln“ (Jungermann, Pfister & Fischer, 2010, S. 3).

Jungermann et al., (2010) beschreiben den Begriff der Entscheidung als einen Problemlöseprozess, in dem eine Person wählt und beurteilt. Der Prozess beginnt zumeist damit, dass die Person feststellt, dass mindestens zwei Optionen bzw. Alternativen zur Auswahl stehen und endet damit, dass die Person sich durch Festlegung auf eine Alternative entscheidet. In einer Entscheidungssituation stehen dem Entscheider demnach mehrere Handlungsoptionen zur Auswahl. Die einzelnen wählbaren Handlungsalternativen führen folglich zu einer bzw. mehrerer Handlungskonsequenzen. Von der Person wird schlussendlich jene Handlungsoption ausgewählt, durch die eine effektive Erreichung eines bestimmten Ziels am optimalsten erscheint (Lerch et al., 2000).

2.4.1.2 Entscheidung unter Risiko

Entscheidungssituationen können grundsätzlich in folgende zwei Bereiche unterteilt werden (Lerch et al., 2000):

- Entscheidungen, die unter Sicherheit getroffen werden,
- Entscheidungen, die unter Unsicherheit bzw. Risiko getroffen werden.

Entscheidungen unter Sicherheit sind beobachtbar, wenn eine Person hinreichende Informationen hinsichtlich der zur Auswahl stehenden Handlungsalternativen und Sicherheit über deren Konsequenzen hat (Kirchler, 1999). Ein Schüler, der beispielsweise völlig unvorbereitet zu einer schriftlichen Prüfung geht, an der die gesamte Klasse teilnimmt, kann mit Sicherheit davon ausgehen, dass seine Unvorbereitetheit zu einem schlechten Prüfungsergebnis führen wird. Bei Entscheidungen, die unter Sicherheit getroffen werden, tritt demnach mit Sicherheit eine bestimmte Folge bzw. Konsequenz der Handlung ein (Lerch et al., 2000).

Bei Entscheidungen, die unter Risiko getroffen werden, führen die zur Auswahl stehenden Handlungsalternativen jedoch nicht mit Sicherheit, sondern nur mit einer bestimmten Wahrscheinlichkeit zu einer bestimmten Konsequenz (Kirchler, 1999; Lerch et al., 2000). Um dies mit einem Beispiel zu verdeutlichen, bleibt es bei einem Schüler, der völlig unvorbereitet zu einer mündlichen Prüfung geht, offen ob nun genau er oder doch ein anderer Mitschüler diese mündliche Prüfung ablegen muss. Die Entscheidung, nicht für die Prüfung zu lernen mit der Folge eine schlechte Beurteilung zu bekommen, tritt nicht mit Sicherheit ein. Die Entscheidung stellt in diesem Fall eine Entscheidung unter Risiko bzw. Unsicherheit dar (Lerch et al., 2000).

Im Folgenden soll nun die SEU- Theorie kurz erörtert werden und auf die Prospect-Theorie, die eine Weiterentwicklung der SEU-Theorie darstellt, etwas genauer eingegangen werden. Gemeinsam ist beiden Theorien,

dass sie einen Erklärungsansatz für Entscheidungen unter Unsicherheit bzw. Risiko bieten.

2.4.1.3 SEU-Theorie- Subjectively Expected Utility Modell

In der SEU-Theorie wird überprüft, ob ein Entscheider gemäß den Kriterien des SEU-Modells Entscheidungen trifft. Nach der SEU-Theorie wird eine Person eine Entscheidung in Hinblick auf die zu wählende Alternative immer nach dem subjektiv höchsten Nutzen, den sie sich von einer bestimmten Alternative erwartet, treffen (Jungermann et al., 2010).

Die Prospect Theorie von Kahnemann und Tversky, entwickelt im Jahr 1979, stellt eine Weiterentwicklung des SEU- Modells dar. Laut Prospect-Theorie sind Menschen darauf bedacht, ausgehend von einem Ausgangspunkt, Gewinn und Verlust einzuschätzen und abzuwägen. Den zu erwartenden Gewinnen oder Verlusten wird jeweils ein subjektiver Wert des Nutzens zugeordnet, der sogenannte Nutzenwert ergibt sich wiederum aus einer subjektiven Bewertungsfunktion (Kahnemann und Tversky, 1979, zitiert nach Fischer & Wiswede, 2009, S.235).

In der Prospect-Theorie wird im Entscheidungsprozess zwischen zwei Phasen und zwei unsicheren Alternativen unterschieden. Vor der Bewertung des Nutzens geht ein Prozess der Bearbeitung der unterschiedlichen Alternativen einher, beispielsweise kommt es zur Selektion von Konsequenzen, die zu identischen Ergebnissen führen. In einem zweiten Prozess werden die bearbeiteten Alternativen nun bewertet, d.h. es wird jeder Alternative ein subjektiver Wert zugeordnet und im Anschluss daran eine Alternative ausgewählt. Gemeinsamkeiten der Prospect- und der SEU-Theorie sind, dass bei beiden davon ausgegangen wird, dass die Konsequenzen und die Wahrscheinlichkeiten einer Alternative die Bewertung der Optionen und damit auch die Richtung der Entscheidung bestimmen. Im Gegensatz zu ihrem Vorläufer geht die Prospect-Theorie von differenziellen Annahmen aus, welche die Umwandlung der Folgen und Wahrscheinlichkeiten in subjektive Größen betreffen (Jungerman et al., 2010). Die Konsequenzen einer Alternative

werden immer in Anbetracht eines bestimmten Referenzpunktes bewertet. Um dies mit einem Beispiel zu veranschaulichen, könnte man annehmen, dass für eine Person, die sich in Lohnverhandlungen mit ihrem Chef befindet und einen Lohn von etwa 3000,- Euro erwartet, ein Lohnvorschlag von 2500,- wenig wert sein. Erwartet sich die Person jedoch einen Lohn von 2000,-, so wird sie über einen Lohnvorschlag von 2500,- wohl sehr erfreut sein. In beiden Fällen bezieht die Person ihre Bewertung von einem Lohnvorschlag über 2500,- auf unterschiedliche Punkte, im ersten Fall auf den 3000,- Euro Betrag und im zweiten Fall auf den 2000,- Euro Betrag. Die gleiche Folge bzw. Konsequenz kann einen unterschiedlichen subjektiven Wert bekommen, je nachdem, von welchem Referenzpunkt man ausgeht. Folgen, die über dem Referenzpunkt liegen, werden als Gewinne angesehen, jene unterhalb als Verluste. Die Wertefunktion kann für Gewinne als konkav und für Verluste als konvex angenommen werden (Jungerman et al., 2010).

2.4.2 RISIKOVERHALTEN UND GESCHLECHT

Den Rahmen einer Vielzahl von Studien bildete die Frage nach Geschlechterunterschieden in Bezug auf Risikoverhalten in den verschiedensten Dimensionen und es konnte immer wieder beobachtet werden, dass die Geschlechter sich in Bezug auf ihr Risikoverhalten unterscheiden (Eckel & Grossmann, 2008).

In einer Untersuchung von Levin, Synder & Chapman (1988) wurden über 100 StudentInnen 18 Glücksspiele vorgegeben, die sich in Bezug auf den Ersteinsatz zu Spielbeginn, die Höhe des Gewinns und die verschiedenen Wahrscheinlichkeiten in Bezug auf den Gewinn unterschieden. Anhand der Ergebnisse konnte ermittelt werden, dass Männer gegenüber Glücksspielen offener als Frauen eingestellt sind.

In einer Untersuchung von Brining (1995) sollten die 300 Testteilnehmer, die sich aus Kindern, Jugendlichen und Erwachsenen zusammensetzten, eine von drei Urnen wählen. In einer dieser drei Urnen versteckte sich ein

„winning ball“. Die erste Urne versprach eine 90-prozentige Chance auf Gewinn, die zweite versprach eine Gewinnchance von 20% mit einem schon etwas größeren Lotteriegewinn und die dritte schließlich versprach eine fünfprozentige Chance auf Gewinn mit einem sehr großen Lotteriegeschenk. Den Ergebnissen zufolge konnten keine Unterschiede der Geschlechter in der Risikoentscheidung gefunden werden. Wurde jedoch das Geschlecht mit dem Alter in Zusammenhang gebracht, so konnte man Unterschiede in Bezug auf das Risikoverhalten beobachten.

In einer Untersuchung von Harbaugh, Krause & Vesterlund (2002) hatten die Versuchsteilnehmer, welche aus Kindern, Teenagern, StudentInnen und Erwachsenen bestanden, 14 Mal die Möglichkeit, zwischen einem sicheren Gewinn und einem Gewinnspiel zu wählen. In dieser Untersuchung konnten keine Geschlechterunterschiede in Bezug auf Risikoverhalten beobachtet werden, auch nicht dann, wenn man das Geschlecht und das Alter miteinander in Beziehung setzte.

In einer Metastudie von Byrnes, Miller und Schafer (1999) wurden die Ergebnisse von insgesamt 150 Studien, in denen Geschlechtsunterschiede im Risikoverhalten untersucht wurden, verglichen. Abhängig von der jeweiligen Situation, in der Risikoverhalten untersucht wurde und dem Alter der Personen kann grundsätzlich aufgrund der Ergebnisse angenommen werden, dass Männer wahrscheinlicher zu riskantem Verhalten neigen als Frauen.

Beispielweise konnten über alle Altersstufen hinweg nur minimale Geschlechtsunterschiede in Bezug auf das Rauchen als riskantes gesundheitsgefährdendes Verhalten festgestellt werden. Zudem waren es ebenso Männer, welche zu riskantem Verhalten neigten, obwohl es in der bestimmten Situation irrational war, sich riskant zu entscheiden. Gegenteilige Ergebnisse konnten für Frauen ermittelt werden, diese verhielten sich auch in Situationen, in denen es möglicherweise angebracht gewesen wäre, riskant zu handeln, weil eventuell die Möglichkeit auf Erfolg steigt, trotzdem risikoaversiv. Zudem waren die

ermittelten Geschlechtsunterschiede bedeutend größer, wenn es um ein real zu beobachtendes Risikoverhalten ging, als in jenen Untersuchungen in denen die VersuchsteilnehmerInnen nur eine mögliche hypothetische Entscheidung treffen sollten (Byrnes et al., 1999).

2.4.2.1 Geschlechterunterschiede im Risikoverhalten im Leistungsbereich mit Multiple Choice Antwortformat

Die Untersuchung von Ben-Shakar & Sinai (1991) beschäftigt sich mit den Geschlechterunterschieden im Zusammenhang mit Risikobereitschaft in einer Testsituation mit Multiple Choice Antwortformat. Im Speziellen galt es zu untersuchen, ob Geschlechterunterschiede in der Tendenz, Items auszulassen, existieren. Da davon ausgegangen werden kann, dass Männer sich risikobereiter zeigen, wird erwartet, dass Männer sich auch in folgender Untersuchung risikofreudiger verhalten als Frauen und somit weniger Items bei der Bearbeitung auslassen. Dies sollte sowohl unabhängig von den Themenbereichen, die der Test erfassen soll, als auch unabhängig von deren Fähigkeiten sein (Ben-Shakar & Sinai, 1991). Zwei Testverfahren, die beide unterschiedliche Subtests zur Erfassung von mathematischen, räumlichen (zwei- und dreidimensional), mechanischen und sprachlichen Fähigkeiten umfassten, wurden eingesetzt. Bei einer der beiden Testformen wurden die Teilnehmer bei der Instruktion aufgefordert zu raten, wenn sie sich hinsichtlich der Lösung der Aufgaben nicht sicher waren. Beim zweiten Test gab es die Instruktion zu raten nicht. Zudem wurden die Teilnehmer darüber informiert, dass es bei Falschantworten keine Strafpunkte gab (Ben-Shakar & Sinai, 1991).

Ziel der Untersuchung war es, den Zusammenhang zwischen der Tendenz zu raten und dem Geschlecht bei der Bearbeitung von Multiple Choice Aufgaben zu ermitteln. Es konnte beobachtet werden, dass Frauen eher dazu neigen, Items auszulassen als ihre männlichen Kollegen. Dies zeigte sich nicht nur in jenen Subtests, bei denen Männer bessere Leistungen zeigen als Frauen, sondern auch in jenen, in denen Frauen gegenüber Männern besser abschnitten, tendierten sie dazu eher, Items

auszulassen als Männer. Im ersten Testverfahren konnten weniger Geschlechterunterschiede in Bezug auf die Ratetendenz ermittelt werden als bei dem zweiten Testverfahren (Hadassah-Test), für die Ursachen werden von den Autoren mehrere Faktoren, wie beispielsweise das Alter der Testteilnehmer, die unterschiedlichen Instruktionen vor Testbeginn oder die Inhalte, die der Test erfasst, genannt.

Zudem konnte ermittelt werden, dass auch, sobald die Teilnehmer vor Testbeginn zum Raten bei Unsicherheit aufgefordert worden waren, die Geschlechterunterschiede hinsichtlich der Tendenz zu raten nicht verschwanden (Ben-Shakar & Sinai, 1991).

2.4.3 GAME OF DICE TASK (GDT)

Der Game of Dice Task (GDT) stellt eine computerisiertes Verfahren dar, welches mittels Würfelspielaufgaben das Entscheidungsverhalten unter Risiko zu ermitteln versucht (Brand et al., 2004).

Mittels Instruktion am Bildschirm werden die Teilnehmer zunächst über die Bedingungen des Spiels genau informiert, die Regeln für Gewinne und Verluste bzw. die Wahrscheinlichkeiten auf Gewinn sind im Vorhinein bekannt (Brand et al., 2004). Zusätzlich sind die Durchgänge des Würfelspiels für die Personen am Computerbildschirm ersichtlich (siehe Abbildung 1).

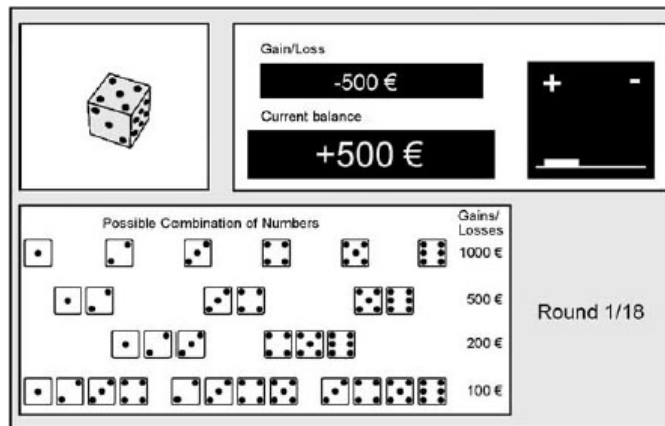


Abbildung 1: Game of Dice Task (GDT). Nach der Entscheidung für eine Zahlenkombination wird gewürfelt und angezeigt, ob der Wurf erfolgreich oder nicht erfolgreich und der Gewinn/ Verlust ist am Bildschirm ersichtlich (Brand et al., 2005)

In einer Studie von Brand et al. (2005) wurde das Entscheidungsverhalten in Risikosituationen mittels des Game of Dice Task (GDT) an einer Stichprobe von 25 männlichen pathologischen Glückspielern und 25 männlichen gesunden Personen untersucht.

Die Ergebnisse belegen, dass die Teilnehmer mit pathologischem Spielverhalten, wenn sie über die Bedingungen der Risikosituation informiert waren, sich hinsichtlich ihres Entscheidungsverhaltens beeinträchtigt zeigten. So trafen sie riskantere Spielentscheidungen als die Teilnehmer der gesunden Kontrollgruppen. Diesbezüglich konnte ein hoher Zusammenhang zwischen dem ungünstigen Entscheidungsverhalten und exekutiven Funktionen, wie beispielsweise Kategorisierung oder kognitiver Flexibilität, festgestellt werden (Brand et al., 2005).

In einer Untersuchung von Euteneuer et al. (2009) wurde an Parkinson Patienten das Entscheidungsverhalten unter Risiko mittels des GDT untersucht. Auch in dieser Untersuchung konnte beobachtet werden, dass die Teilnehmer mit einer Parkinson-Erkrankung ein deutlich ungünstigeres Entscheidungsverhalten zeigten als die gesunde Kontrollgruppe. Des Weiteren konnte ein bedeutender Zusammenhang zwischen Exekutiven Dysfunktionen und dem GDT gezeigt werden. Es kann angenommen

werden, dass ein Zusammenhang zwischen dem Entscheidungsverhalten im GDT und Dysfunktionen im dorsolateralen Präfrontalkortex, welche als Resultat der Defizite exekutiver Funktionen verstanden werden können und wiederum eine wesentliche Rolle beim Entscheidungsprozess spielen, besteht (Euteneuer et al., 2009).

3 EMPIRISCHER TEIL

3.1 ZIELSETZUNG, FRAGESTELLUNG UND HYPOTHESEN

Das Ziel dieser empirischen Arbeit ist die Untersuchung von Geschlechterunterschieden im Rahmen der Bearbeitung von EMS-ähnlichen Aufgaben. Von besonderem Interesse sind jene Bereiche bzw. Untertests der EMS-ähnlichen Aufgaben, welche die Merkfähigkeit erfassen. Des Weiteren soll erforscht werden ob neben dem Geschlecht noch andere soziale Faktoren wie Stereotype Threat oder Risikoverhalten einen bedeutenden Einfluss auf die kognitiven Leistungsfähigkeit haben.

Zudem wird der Einfluss von Persönlichkeitsvariablen, wie Selbstwirksamkeitserwartung und Leistungsmotivation erhoben, deren detaillierte Ergebnisse in den Arbeiten von Gabriele Hangl und Anita Teufl nachzulesen sind. Ebenso werden Geschlechtsunterschiede in Bezug auf die Theory of Mind, als auch Unterschiede zwischen Männern und Frauen in EMS-ähnlichen Aufgaben untersucht, welche das Schlussfolgernde Denken betreffen. Detaillierte Ergebnisse hierzu finden sich in den Arbeiten von Katrin Anzirk und Silvia Hameseder.

Zu Beginn des jeweiligen Abschnitts und vor der Überprüfung der statistischen Hypothesen werden die zu untersuchenden und interessierenden Hypothesen in Form eines Hypothesenblocks dargestellt.

3.2 METHODE

Die vorliegende Diplomarbeit stellt den Teil eines größer angelegten Forschungsprojektes dar und wurde 2009 von der medizinischen Universität Wien in Auftrag gegeben.

Hauptziel der Untersuchung ist es, Geschlechterunterschiede und Einflussfaktoren in Bezug auf das Medizinaufnahmeverfahren EMS zu

ermitteln, um in weiterer Folge Anhaltspunkte und Hinweise für eine faire Gestaltung des Eingangstests zu erlangen.

Die KollegInnen Katrin Anzirk, Silvia Hameseder, Gabriele Hangl und Anita Teufl beschäftigen sich in ihren Diplomarbeiten mit speziellen Schwerpunkten und spezifischen Aspekten dieses Forschungsprojektes. Im Detail mit der Theory of Mind, den EMS- ähnlichen Untertests zum schlussfolgernden Denken, der Leistungsmotivation und der Selbstwirksamkeitserwartung.

Im folgenden Kapitel werden nun Untersuchungsplan, intendierte Stichprobe und die tatsächliche Untersuchungsdurchführung näher beschreiben. Zudem erfolgt eine Darstellung der Erhebungsinstrumente, welche in der Untersuchung eingesetzt wurden.

3.2.1 UNTERSUCHUNGSPLAN UND INTENDIERTE STICHPROBE

Das ursprüngliche Ziel des Forschungsprojektes war ein multivariater Ansatz um vorhandene geschlechtsspezifische Unterschiede in kognitiven Leistungssituationen mittels EMS-ähnlichen Aufgaben unter Berücksichtigung von biologischen Faktoren, wie beispielsweise dem Hormonstatus und sozialen Faktoren, wie Stereotype Threat und Risikobereitschaft und der Vorgabe von Theory of Mind Aufgaben aufzudecken. Desweiteren galt es etwaige Mediatorvariablen wie Ängstlichkeit, Leistungsmotivation und Selbstwirksamkeit zu erfassen.

Methodisch war ein experimentelles multivariates Design mit zufälliger Zuteilung der UntersuchungsteilnehmerInnen in die Versuchsbedingungen (Stereotype Threat und Risikoverhalten) geplant. Die Untersuchung besteht insgesamt aus 24 vollständigen Kombinationen der Faktorstufen, die sich aus dem Geschlecht (2), Stereotype Threat (4) und Risikoverhalten (3) zusammensetzen.

In der folgenden Tabelle sollen die experimentellen Versuchsbedingungen (Stereotype Threat und Risikoverhalten) graphisch veranschaulicht werden.

Tabelle 2: Studiendesign und Zuweisung in die Versuchsbedingungen

<u>Variablen</u>	<u>Versuchsbedingungen</u>											
Hormonmessung												
Stereotype Threat (4)	<u>A</u>			<u>B</u>			<u>C</u>			<u>D</u>		
Risikoverhalten (3)	1	2	3	1	2	3	1	2	3	1	2	3
Geschlecht (m/w)												
EMS-ähnlicher Test												
Theory of Mind												
Hormonmessung												
Ängstlichkeit/ Selbstkonzept & Leistungsmotivation												

Um die Repräsentativität der Untersuchung zu gewährleisten, war geplant, eine möglichst große Stichprobe (etwa 1200 Testpersonen) zu rekrutieren. Für die Untersuchung sollten SchülerInnen, welche die zwölfte Schulstufe einer Allgemein Bildenden Höheren Schule (AHS) besuchen. Die UntersuchungsteilnehmerInnen sollten an AHS in Wien und Niederösterreich rekrutiert werden.

Neben der Vorgabe eines Soziodemographischen Fragebogens zu Beginn der Testung soll im Anschluss an die Vorgabe der EMS-ähnlichen Aufgaben, im Rahmen von 6 Untertests mit zeitlicher Beschränkung, Items zu Theory of Mind vorgegeben werden. Zusätzlich sollte vor und nach der Testung eine Speichelprobe von den SchülerInnen entnommen werden um eine Hormonuntersuchung zu ermöglichen.

Da es sich bei der Stichprobe um SchülerInnen der Allgemein Bildenden Höheren Schulen (AHS) in Wien und Niederösterreich handeln soll, war geplant vor Beginn der Untersuchung einen Antrag für die Genehmigung der Durchführung dieser Studie beim jeweils zuständigen Stadt- bzw. Landesschulrat und zudem einen Antrag bei der Ethikkommission der Medizinischen Universität Wien zu stellen.

Noch vor dem Start der eigentlichen Erhebungsphase war eine Voruntersuchung geplant, um die praktische Umsetzung und den zeitlichen Rahmen der Untersuchung zu ermitteln.

Geplanter Beginn der Untersuchung war Oktober 2009 im Anschluss Erhebungsphase sollte die Dateneingabe und Auswertung folgen, für die ein Zeitraum von etwa drei Monaten eingerechnet war und schlussendlich sollte es im Juni 2010 zum Abschluss der Untersuchung kommen.

3.2.2 ERHEBUNGSINSTRUMENTE

Im Folgenden Teil werden alle Instrumente beschrieben, die im Rahmen der Untersuchung eingesetzt wurden. Zuvor sollen die während der Untersuchung vorgegebenen Stereotype Threat- Bedingungen und Risikoinstruktionen kurz beschrieben und erklärt werden.

Insgesamt zur Anwendung kamen folgende Messinstrumente:

- ❖ Soziodemographischer Fragebogen
- ❖ Wissenstest
- ❖ EMS-ähnliche Aufgaben (ITB Consulting, 2008; Zentrum für Testentwicklung und Diagnostik [ZTD], 2009); Schlauchfiguren (Stumpf & Fay, 1983)
- ❖ Theory of Mind Aufgaben (Willinger, Schmöger, Müller & Auff, in Vorbereitung)
- ❖ Leistungsmotivationsinventar (Schuler & Prochaska, 2001), Skalen: (1) Wettbewerbsorientierung, (2) Internalität
- ❖ Allgemeine Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999)

❖ Game of Dice Task (Brand et al., 2005)

3.2.2.1 Experimentelle Versuchsbedingungen

Vor jedem der 6 EMS-ähnlichen Untertest werden den UntersuchungsteilnehmerInnen 4 verschiedene Stereotype Threat Bedingungen und 3 verschiedene Risikoinstruktionen vorgegeben. Im Anschluss sollen diese genauer erklärt werden.

Stereotype Threat- Bedingung

- (1) In der *Aktivierung von Stereotype Threat* wurden die UntersuchungsteilnehmerInnen dahingehend schriftlich instruiert, es sei laut Literatur bekannt, dass Frauen in kognitiven Aufgaben, welche das Gedächtnis betreffen besser abschneiden und Männer hingegen in kognitiven Aufgaben die das Arbeiten mit Zahlen, die räumliche Orientierung, das Verstehen von komplexen Texten und das naturwissenschaftliche Verständnis betreffen besser abschneiden.
- (2) In der *Neutralen Stereotype Threat* Bedingung wurde den TeilnehmerInnen mitgeteilt, dass es keine geschlechtsspezifischen Unterschiede im Bearbeiten von kognitiven Aufgaben gibt.
- (3) In der Bedingung *Stereotype Threat Aufklärung* wurde den Versuchspersonen das Phänomen Stereotype Threat erklärt.
- (4) Anders als in der Bedingung der Aktivierung von Stereotype Threat wurde den TestteilnehmerInnen in der *Negativaktivierung von Stereotype Threat* eine genau gegenteilige Information bezüglich Geschlechtsunterschieden hinsichtlich kognitiver Leistungsfähigkeit mitgeteilt, nämlich dass Frauen bessere Leistungen im Bearbeiten von Aufgaben , die räumliche Orientierung, das Verstehen von komplexen Texten, Arbeiten mit Zahlen und naturwissenschaftliches Verständnis betreffen und Männer bessere Leistungen im Gedächtnisaufgaben zeigen.

Risikoverhalten

- (1) In der ersten Bedingung, der *Ratebedingungen* wurden die TestteilnehmerInnen aufgefordert zu Raten auch wenn sie sich hinsichtlich der Lösung der Aufgaben nicht sicher waren.
- (2) In der zweiten Bedingung, der Bedingung *ohne spezielle Instruktion* wurde den TestteilnehmerInnen nur gesagt, sie sollen die Aufgaben aufmerksam durchlesen und in weiterer Folge beantworten.
- (3) In der dritten Risikobedingung teilte man den TeilnehmerInnen mit es sollen nur die Aufgaben beantwortet werden, von denen man glaubt, dass man sie *sicher richtig beantworten* kann.

3.2.2.2 Soziodemographischer Fragebogen

Mit dem soziodemographischen Fragebogenteil, der zu Beginn der Testung vorgegeben wurde, sollten Personenbezogene Information wie beispielsweise Alter, Geschlecht und Nationalität der TestteilnehmerInnen erhoben werden. Desweiteren sollten Informationen bezüglich des besuchten Schulzweiges und Beruf der Eltern ermittelt werden. Ein weiterer Teil des Fragebogens bezieht sich speziell auf Fragen, die den weiteren Bildungsweg betreffen, ob der Wunsch auf ein Medizinstudium vorhanden ist und wenn ja ob man sich schon mit der Vorbereitung des Medizinaufnahmeverfahrens beschäftigt hat.

3.2.2.3 Wissenstest

Der sogenannte Wissenstest wurde im Anschluss an den Soziodemographischen Fragebogen ohne zeitliche Beschränkung vorgegeben. Mit dem Wissenstest sollte überprüft werden ob die zuvor instruierten Stereotype Threat Bedingungen ausgelöst wurden und basiert auf der Studie von Halpern und Tan (2001), *Beliefs About the Cognitive Abilities of Females and Males Questionnaire*. Der sogenannte Wissenstest besteht aus insgesamt 14 Aufgaben. Im Ersten Teil, bestehend aus 7 Items, zu jeweils einem Themenbereich der 6 EMS-ähnlichen Aufgaben und einer Theory of Mind Aufgabe, sollen die

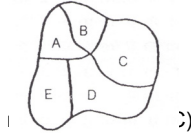
TestteilnehmerInnen sich eine Person, die sie nicht kennen vorstellen und in Prozent einschätzen ob diese Aufgabe wahrscheinlicher von einer Frau bzw. einem Mann gelöst wurde. Die Summe der Einschätzung sollte immer 100% ergeben (siehe Abbildung 2).

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

Einprägen der schwarzen Flächen verschiedener Figuren:



Welche Fläche war die geschwärzte?



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 2: Beispielaufgabe Wissenstest Figuren Lernen

In weiteren 7 Aufgaben wurden die TestteilnehmerInnen aufgefordert zu wiederum 6 EMS- Themenbereiche und einer Theory of Mind- Aufgabe sich selbst auf einer siebenstufigen Skala einzuschätzen ob sie glauben die Fähigkeiten zu besitzen die entsprechende Aufgabe richtig lösen zu können (siehe Abbildung 3).

Denken sie an folgende Aufgabe: Eine Person muss sich Fakten einprägen, es wurden ihr 15 PatientInnen vorgestellt. Sie erfährt jeweils den Namen, die Altersgruppe, Beruf und Geschlecht, ein weiteres Beschreibungsmerkmal (z.B. Familienstand) sowie die Diagnose der PatientInnen. Bei einer späteren Befragung musste die Person bestimmte Details wiedergeben, beispielsweise „der Patient mit dem Schädelbasisbruch ist von Beruf...

A) Installateur
B) Lehrer
C) Dachdecker
D) Handelsvertreter
E) Physiker

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: Beispielaufgabe Wissenstest Fakten Lernen

3.2.2.4 EMS-ähnliche Aufgaben

Aus den insgesamt zehn Untertest des EMS, welche jährlich für das Medizinaufnahmeverfahren in Deutschland, der Schweiz und Österreich zum Einsatz kommen, sind im Rahmen dieser Untersuchung sechs Untertests für die Vorgabe ausgewählt worden. Einerseits aus zeittechnischen Gründen und andererseits fiel die Wahl im speziellen auf jene sechs Untertests, welche Bereiche erfassen, deren Ergebnisse in der Vergangenheit entweder zugunsten der Frauen bzw. zugunsten der Männer ausfielen. Insgesamt wurden absichtlich mehr Aufgaben pro Untertest vorgegeben, als innerhalb der vorgegebenen Zeit schaffbar gewesen wären und werden im Multiple Choice Antwortformat vorgegeben.

Im Folgenden sind die sechs Untertests der EMS-ähnlichen Aufgaben genauer beschrieben:

Untertest Fakten Lernen und Figuren Lernen

Die Untertests Fakten und Figuren Lernen überprüfen die Gedächtnisleistungen der TestteilnehmerInnen und zeichnen sich durch ihre analogen Vorgabemodi aus.

Untertest Fakten Lernen

In diesem Untertest soll überprüft werden wie gut die TestteilnehmerInnen sich Fakten (vorwiegend medizinischer Inhalt) einprägen können und zu einem späteren Zeitpunkt der Testdurchführung wieder abrufen können.

Der Untertest Fakten Lernen besteht aus einer *Einprägephase* und einer *Reproduktionsphase*. Für das Einprägen der Fakten sind drei Minuten vorgesehen, diese soll die Testperson einige Untertests später wieder reproduzieren können, hierfür sind ebenso drei Minuten vorgesehen. Insgesamt wurden acht Items vorgegeben.

In der Einprägephase soll sich die Person Fakten zu 8 Patienten einprägen, zu welchen sie in der Reproduktionsphase Frage beantworten soll.

Holz: ca. 35 Jahre, Heilpädagogin, Krankmeldung- Magengeschwür

Die Diagnose für die Heilpädagogin lautet:

- (A) Bluthochdruck
- (B) Kreislaufkollaps
- (C) Herzinfarkt
- (D) Karies
- (E) Magengeschwür

Abbildung 4: Beispielaufgabe EMS- ähnlicher Untertest Fakten Lernen Einpräge- und Reproduktionsphase (ZTD, 2009)

Untertest Figuren Lernen

Der Untertest Figuren Lernen erfasst wie gut sich die TestteilnehmerInnen Einzelheiten, welche visuell dargeboten werden merken können, diese behalten und zu einem späteren Zeitpunkt wieder abrufen können.

Analog zum Untertest Fakten Lernen besteht auch der Untertest Figuren Lernen aus einer *Einpräge-* und einer *Reproduktionsphase*. Für das Einprägen wie auch für das Reproduzieren der acht Items hat die Testperson zwei Minuten Zeit.

In der Einprägephase soll sich die Person die schwarzen Flächen einprägen (siehe Abbildung). In der Reproduktionsphase soll angegeben werden welcher Teil der Abbildung geschwärzt war (siehe Abbildung).



Abbildung 5: Beispielaufgabe EMS- ähnlicher Untertest Figuren Lernen Einpräge- und Reproduktionsphase (ZTD, 2009)

Untertest Quantitative und Formale Probleme

Der Untertest Quantitative und Formale Probleme erfasst Fähigkeiten im Umgang mit Zahlen, Einheiten und Formeln in einem medizinisch bzw. naturwissenschaftlichen Kontext.

Insgesamt wurden zu diesem Untertest 16 Items vorgegeben und dafür haben die TestteilnehmerInnen 10 Minuten Zeit.

Ein physikalisches Gesetz ermöglicht die Bestimmung einer Kraft F aus den Größen

Geschwindigkeit	v	(Einheit: 1 m/s),
Masse	m	(Einheit: 1 kg),
Radius	r	(Einheit: 1 m).

Bei Anwendung welcher der folgenden fünf Formeln ergibt sich für F die Einheit $1 \text{ kg} \cdot \text{m/s}^2$?

(A) $F = v^2 \cdot r / m$
(B) $F = r \cdot \sqrt{v \cdot m}$
(C) $F = m \cdot v^2 / r$
(D) $F = m^2 \cdot v \cdot r$
(E) $F = r \cdot v^2 \cdot m$

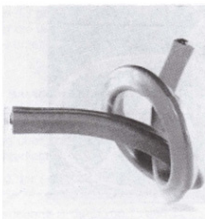
Abbildung 6: Beispielaufgabe EMS- ähnlicher Untertest Quantitative und Formale Probleme (ZTD, 2009)

Untertest Schlauchfiguren

Die Aufgaben des Untertests Schlauchfiguren erfassen die Fähigkeit des räumlichen Vorstellungsvermögens.

Insgesamt wurden den Testpersonen 21 Aufgaben zur Bearbeitung vorgegeben und für diese hatten sie 12 Minuten Zeit.

Die erste Abbildung (links) zeigt die Vorderansicht des Würfels. Die zweite Abbildung (rechts) zeigt denselben Würfel noch einmal. Die Person soll angeben ob der Würfel auf der rechten Abbildung von links, rechts, unten, oben oder hinten dargestellt ist.



(A): r
(B): l
(C): u
(D): o
(E): h

Abbildung 7: Beispielaufgabe EMS- ähnlicher Untertest Schlauchfiguren (ZTD, 2009)

Untertest Medizinisch-naturwissenschaftliches Grundverständnis

Aufgaben, die im Untertest Medizinisch-naturwissenschaftliches Grundverständnis vorgelegt werden, erfassen ein grundlegendes Verständnis für Medizin und Naturwissenschaften.

Dieser Untertest umfasst 16 Aufgaben mit einer Bearbeitungszeit von 10 Minuten.

Im menschlichen Körper ist die sog. Stickstoffbilanz normalerweise ausgeglichen, d.h. die Menge des aufgenommenen Stickstoffs entspricht der des ausgeschiedenen. Der aufgenommene Stickstoff ist in den Eiweißen der Nahrung enthalten. Wird mehr Stickstoff aufgenommen als über die Nieren ausgeschieden, spricht man von einer positiven Stickstoffbilanz, im umgekehrten Fall von einer negativen Stickstoffbilanz.

Im Hungerzustand werden körpereigene Eiweiße abgebaut und als Energielieferanten verwendet. Dabei werden ihre Bausteine, die Aminosäuren, aufgespalten, und der anfallende Stickstoff wird im Harn ausgeschieden.

Wie sieht die Stickstoffbilanz im Hungerzustand aus?

- (A) Sie ist positiv, da mehr Stickstoff ausgeschieden als aufgenommen wird.
- (B) Sie ist positiv, da mehr Stickstoff ausgeschieden wird als normalerweise.
- (C) Sie ist negativ, da weniger Stickstoff ausgeschieden wird als normalerweise.
- (D) Sie ist ausgeglichen, da im Hungerzustand entsprechend weniger Stickstoff ausgeschieden wird.
- (E) Sie ist negativ, da mehr Stickstoff ausgeschieden wird als normalerweise.

Abbildung 8: Beispielaufgabe EMS- ähnlicher Untertest Medizinisch-naturwissenschaftliches Grundverständnis (ZTD, 2009)

Untertest Diagramme und Tabellen

Diese Gruppe von Aufgaben ermittelt die Fähigkeit der Analyse von Diagrammen und Tabellen.

Die TestteilnehmerInnen haben auch für diesen Themenblock 10 Minuten Zeit zur Bewältigung von 16 Aufgaben.

Die folgende Tabelle beschreibt die Zusammensetzung und den Energiegehalt von vier verschiedenen Milcharten. Unter Energiegehalt der Milch verstehen wir dabei die Energiemenge, gemessen in Kilojoule (kJ), welche 100 Gramm (g) der Milch dem Organismus ihres Konsumenten liefern können.

Milchart	Ei- weiss	Fett	Milch- zucker	Salze	Energie- gehalt
menschliche Muttermilch	1,2 g	4,0 g	7,0 g	0,25 g	294 kJ
Vollmilch	3,5 g	3,5 g	4,5 g	0,75 g	273 kJ
Magermilch	3,3 g	0,5 g	4,5 g	0,75 g	160 kJ
Buttermilch	3,0 g	0,5 g	3,0 g	0,55 g	110 kJ

Welche Aussage lässt sich aus den gegebenen Informationen nicht ableiten?

- (A) Menschliche Muttermilch enthält mehr als doppelt soviel Fett und mehr als doppelt soviel Milchzucker wie Buttermilch.
- (B) Der Eiweissgehalt der Milch ist für ihren Energiegehalt von entscheidender Bedeutung.
- (C) Zur Aufnahme der gleichen Energiemenge muss ein Säugling fast dreimal soviel Buttermilch wie Muttermilch trinken.
- (D) Der Unterschied zwischen Magermilch und Vollmilch ist bei der Mehrzahl der aufgeführten Merkmale geringer als der Unterschied zwischen Magermilch und Buttermilch.
- (E) Vollmilch enthält im Vergleich zur menschlichen Muttermilch etwa die dreifache Menge an Salzen und Eiweiss.

Abbildung 9: Beispielaufgabe EMS- ähnlicher Untertest Tabellen und Diagramme (ZTD, 2009)

Für Details der erfassten kognitiven Fähigkeit der EMS-Untertests wird der Bericht über den EMS von Mallinger et al., (2009) verwendet.

3.2.2.5 Theory of mind

Nach der Bearbeitung der EMS- ähnlichen Aufgaben wurden den SchülerInnen zur Erfassung der *Theory of Mind* (Willinger et al., in Vorbereitung) ohne zeitlicher Beschränkung drei Geschichten vorgegeben, zu denen im Anschluss Fragen zu beantworten waren.

„Eine *Theory of Mind*, Theorie des Geistes oder des Mentalen, ist ein Konstrukt aus zusammenhängenden Überzeugungen und Wünschen, die man einer Person zuschreibt, um sich deren Verhalten zu erklären.“ (Fonagy, Gergely, Jurist & Target, 2008, S. 34).

3.2.2.6 Leistungsmotivation

Das Leistungsmotivationsinventar von Schuler und Prohaska (2001) ist ein Persönlichkeitsfragebogen, der mit insgesamt 170 Items und 17 Dimensionen vor allem die berufliche Leistungsmotivation erfasst.

Im Rahmen dieser Studie wurden den VersuchsteilnehmerInnen auf einer siebenstufigen Ratingskala zwei Skalen des Leistungsmotivationsinventars, *Wettbewerbsorientierung* und *Internalität* ohne zeitliche Beschränkung vorgegeben, welche insgesamt 20 Items umfassten.

	trifft gar nicht zu	trifft vollständig zu
Wie weit man es beruflich bringt, ist zu einem guten Teil Glückssache.	①-②-③-④-⑤-⑥-⑦	
Es macht mir nichts aus, wenn andere im gleichen Alter mehr erreicht haben als ich.	①-②-③-④-⑤-⑥-⑦	
Manchmal habe ich das Gefühl, dass man es mir schwer macht und mich entmutigt.	①-②-③-④-⑤-⑥-⑦	

Abbildung 10: Beispielaufgaben Leistungsmotivation (Schuler & Prohaska, 2001)

3.2.2.7 Selbstwirksamkeitserwartung

Im Anschluss an die Bearbeitung der *Theory of Mind* Aufgaben und des Leistungsmotivationsinventars wurden den SchülerInnen 10 Items der *Skala der Allgemeinen Selbstwirksamkeitserwartung* von Schwarzer und Jerusalem (1999) vorgelegt, um die Selbstwirksamkeitserwartung der TeilnehmerInnen zu erfassen.

Mit der Skala der allgemeinen Selbstwirksamkeitserwartung wird versucht, Fragen wie beispielsweise die Einschätzung der eigenen Kompetenzen oder das Bewältigen von Schwierigkeiten im Alltag zu erfassen (Schwarzer & Jerusalem, 1999).

Nach Schwarzer und Jerusalem (1999) haben Selbstwirksamkeitserwartungen Einfluss auf die Auswahl von Handlungen und die Erreichung eines Ziels und bestimmen auch, wenn auch nur indirekt den Erfolg einer Handlung.

	stimmt nicht	stimmt kaum	stimmt eher	stimmt genau
Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.	1	2	3	4
Die Lösung schwieriger Probleme gelingt mir immer, wenn ich mich darum bemühe.	1	2	3	4
Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen.	1	2	3	4

Abbildung 11: Beispielaufgaben Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999)

3.2.2.8 Game of Dice Task (GDT)

Der Game of Dice Task (GDT) wurde von der Forschergruppe um Brand (2005) entwickelt. Der GDT ist ein Verfahren zur Messung des Risikoverhaltens. Die TestteilnehmerInnen werden dazu aufgefordert am Computer Würfelspielaufgaben zu spielen, die insgesamt 18 Durchgänge umfassen. Gewinne und Verluste sind den TeilnehmerInnen bekannt und am Computerbildschirm ersichtlich (siehe Abbildung 12).

In der vorliegenden Untersuchung wurde den SchülerInnen der GDT im Anschluss an die EMS-ähnlichen Aufgaben vorgegeben, um zusätzlich zu den Risikoinstruktionen das Risikoverhalten zu erfassen. Die Bearbeitungszeit für den GDT umfasste fünf Minuten.

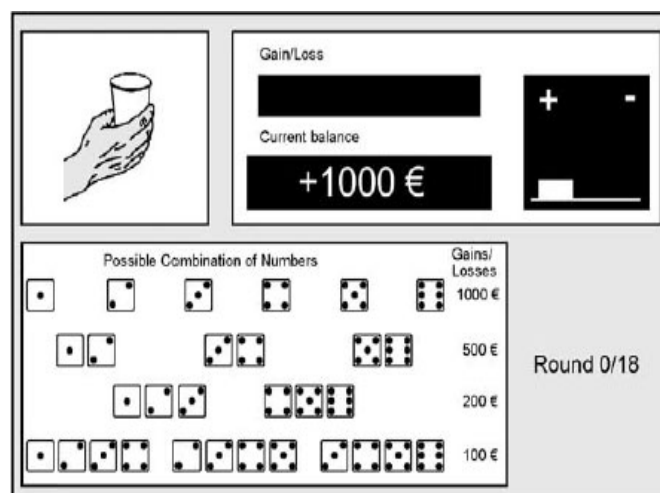


Abbildung 12: Game of Dice Task (Brand et al., 2005)

Zu Beginn jedes Durchgangs sollen sich die Personen zwischen einer Zahl (z.B. 5), einer Kombination aus zwei Zahlen (z.B. 4 und 5), zwischen einer Kombination aus drei Zahlen (z.B. 4,5 und 6) oder vier Zahlen (z.B. 3,4,5 und 6) entscheiden. Danach wird gewürfelt, das Ergebnis wird sichtbar und Gewinn bzw. Verlust werden am Bildschirm angezeigt (Brand et al., 2005).

3.3 UNTERSUCHUNG

3.3.1 UNTERSUCHUNGSDURCHFÜHRUNG

Die vorliegende Untersuchung wurde von der medizinischen Universität in Auftrag gegeben und von Ao. Univ.- Prof. Mag. Dr. Ulrike Willinger betreut. Nach der Erstellung der Testbatterie und Planung der Untersuchung wurde das Konzept zur Genehmigung bei der Ethikkommission der Medizinischen Universität Wien eingereicht und etwa ein Monat später bewilligt.

Um Vorab den zeitlichen Rahmen der Untersuchung und die Durchführbarkeit zu ermitteln wurde eine Voruntersuchung durchgeführt. Des Weiteren sollte überprüft werden ob die Aktivierung der experimentellen Versuchsbedingungen (Stereotype Threat und Risikoinstruktion) umgesetzt und ausgelöst werden konnte. Die Voruntersuchung wurde an StudentInnen der medizinischen Universität Wien durchgeführt, die sich alle am Ende des 2. Abschnitts ihres Studiums befanden. Aufgrund der Ergebnisse der Vorstudie wurde auf die Vorgabe des Untertests *Textverständnis* aus zeitlichen Aspekten verzichtet.

Im Anschluss der Einholung der Genehmigung für die Testung an Allgemeinbildenden Höheren Schulen durch die zuständigen Stadt- bzw. Landesschulräte begann die Phase der Terminvereinbarung- bzw. Koordination mit den Schulen bzw. DirektorInnen. Die Phase der

Rekrutierung der TestteilnehmerInnen stellte sich als besonders zeitaufwendig heraus.

Die Testungen fanden in den teilnehmenden Schulen statt, mit dem zeitlichen Ausmaß von zwei Unterrichtsstunden. Die Hormonuntersuchung wurde mittels Speichelproben vor und nach der Vorgabe der Testbatterie erhoben, konnte aber aus finanziellen Gründen nur an einem Teil der getesteten SchülerInnen ermittelt werden. Aus zeitökonomischen Gründen war die Vorgabe des *Game of Dice Task* (GDT) mittels Computer im Anschluss an die EMS- ähnlichen Aufgaben nur bei etwa der Hälfte der Testpersonen möglich.

Schließlich konnte im März 2010 die Erhebungsphase beginnen und im Februar 2011 abgeschlossen werden.

3.3.2 BESCHREIBUNG DER STICHPROBE

Im folgenden Abschnitt wird die Stichprobe anhand von bedeutsamen Merkmalen, wie Geschlecht, Lebensalter, Nationalität, den besuchten Schulzweigen, dem Wunsch, Medizin zu studieren und dem Wiederholen einer Schulstufe dargestellt.

Des Weiteren wird die Zusammensetzung der Stichprobe in den einzelnen Versuchsbedingungen aufgezeigt und schließlich die Verteilung der relativen Lösungswahrscheinlichkeiten der EMS-ähnlichen Aufgaben dargestellt.

Zur deskriptiven und inferenzstatistischen Datenanalyse wird SPSS (Statistical Package for Social Science), Version 19.0.1, herangezogen.

Verteilung der Geschlechter

Insgesamt konnten 1093 SchülerInnen (N=1093) an Gymnasien in Niederösterreich (20,0 %), Burgenland (7,3 %) und Wien (72,7 %) für die Untersuchung herangezogen werden. Von den insgesamt 1093 SchülerInnen waren 684 (62,6 %) weiblich und 409 (37,4 %) männlich (siehe Tabelle 3).

Tabelle 3: Verteilung der Geschlechter über die Gesamtstichprobe

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

Aus folgender Tabelle ist ersichtlich, dass die Prüfung auf Gleichverteilung mit $\chi^2(2) = 69.190$, $p \leq .0001$ ein signifikantes Ergebnis ergibt und somit ein Verteilungsunterschied im Geschlecht über die Gesamtstichprobe beobachtet werden kann.

Tabelle 4: Chi- Quadrat Test Geschlechterverteilung über die Gesamtstichprobe

Chi-Quadrat	69,190 ^a
df	1
Asymptotische Signifikanz	$p \leq .0001$

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 546,5.

Zur Veranschaulichung wird die Verteilung der Geschlechter über die gesamte Stichprobe graphisch in Abbildung 13 dargestellt.

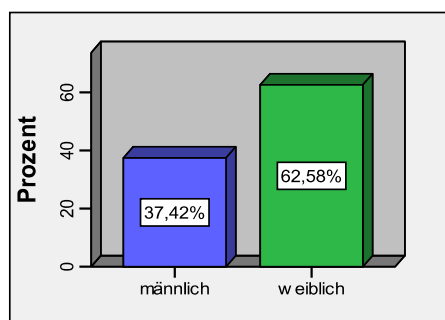


Abbildung 13: Anteilswerte der Geschlechter in Prozent

3.3.2.1 Nationalität der SchülerInnen

Die folgende Abbildung soll einen Überblick über die Verteilung der Nationalitäten der teilgenommenen SchülerInnen geben. Von insgesamt 1050 gültigen Fällen haben 992 (94,48 %) SchülerInnen die österreichische Staatsbürgerschaft angegeben. Weitere 58 (5,52 %) SchülerInnen gaben eine andere Nationalitätszugehörigkeit an, wie beispielsweise Deutschland, Kroatien, Rumänien, Slowakei, Ungarn, Polen, Schweden, Serbien, Japan und Türkei an. Es gab 43 (3,98 %) SchülerInnen von den insgesamt 1093 TeilnehmerInnen (N=1093), welche keine Angaben in Bezug zu ihrer Nationalität machten (siehe Abbildung 14).

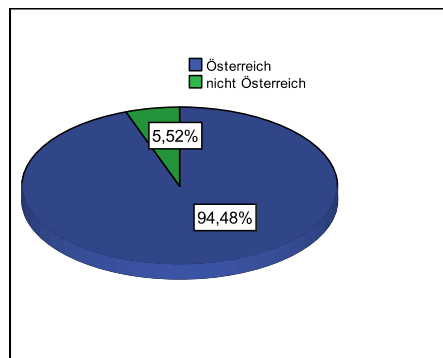


Abbildung 14: Anteilswerte der Nationalität in Prozent

3.3.2.2 Lebensalter der Gesamtstichprobe

Folgende Tabelle gibt einen Überblick über die Häufigkeiten in Bezug auf das Lebensalter der Gesamtstichprobe.

Tabelle 5: Häufigkeiten für Lebensalter über die Gesamtstichprobe

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 Berechnung der entsprechenden Prüfgröße $\chi^2(5) = 1643,116$, $p \leq .0001$ ein signifikantes Ergebnis, somit können Verteilungsunterschiede im Lebensalter beobachtet werden (siehe Tabelle 6).

Tabelle 6: Chi- Quadrat Test für das Lebensalter

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

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 181,0.

Zum Zeitpunkt der Testung befanden sich die SchülerInnen im Alter zwischen 16 und 21 Jahren ($M = 17.57$; $SD = 0.76$). Jene 580 (53,4 %) SchülerInnen, welche zum Testzeitpunkt 17 Jahre alt waren, stellten die größte Kategorie der 1086 TeilnehmerInnen ($N=1086$) dar. Von den insgesamt 1093 TestteilnehmerInnen ($N= 1093$) liegt bei sieben Personen das Alter nicht vor.

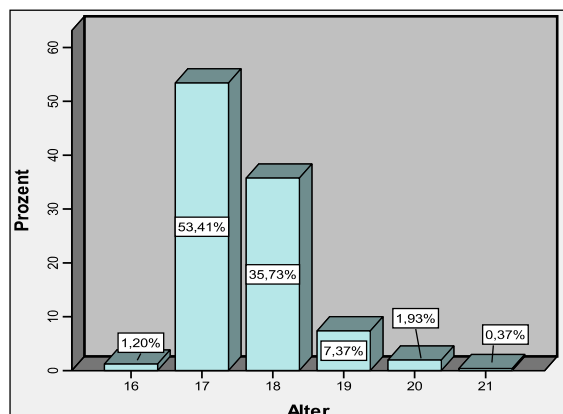


Abbildung 15: Lebensalter der UntersuchungsteilnehmerInnen (1086 gültige Fälle)

3.3.2.3 Schulzweige der SchülerInnen

Tabelle 7 im Anschluss zeigt die Häufigkeiten der besuchten Schulzweige der untersuchten Stichprobe.

Tabelle 7: Häufigkeiten für besuchte Schulzweige

Schulzweige			
	Beobachtetes N	Erwartete Anzahl	Residuum
Naturwissenschaftlich	247	90,6	156,4
Sprachlich	288	90,6	197,4
DG	46	90,6	-44,6
Bildnerisch	67	90,6	-23,6
Musikalisch	199	90,6	108,4
Technisch	1	90,6	-89,6
Wirtschaftlich	27	90,6	-63,6
Humanistisch	20	90,6	-70,6
Informatik	20	90,6	-70,6
Sonstige	68	90,6	-22,6
Sportlich	97	90,6	6,4
fehlend	7	90,6	-83,6
Gesamt	1087		

Die Berechnung der entsprechenden Prüfgröße $\chi^2(11) = 1184.621$, $p \leq .0001$ ein signifikantes Ergebnis. Es können Verteilungsunterschiede in den besuchten Schulzweigen angenommen werden (siehe Tabelle 8).

Tabelle 8: Chi- Quadrat Test für besuchte Schulzweige

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

a. Bei 0 Zellen (.0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 90,6.

Die nachfolgende Tabelle veranschaulicht, welche Schulzweige von den untersuchten TeilnehmerInnen in der AHS-Oberstufe besucht wurden. Jene drei Schulzweige, welche von den 1087 SchülerInnen, von denen Angaben bezüglich ihres besuchten Schulzweiges gemacht wurden, am häufigsten besucht werden, sind mit 288 SchülerInnen (26,5 %) der sprachliche Schulzweig, gefolgt vom naturwissenschaftlichen Schulzweig mit 247 (22,7 %) SchülerInnen und mit 199 (18,2 %) der musikalische Schulzweig. Die restlichen SchülerInnen gaben schwerpunktmäßig einen sportlichen (8,9 %), einen bildnerischen (6,2 %), als auch einen wirtschaftlichen (2,5 %), humanistischen (1,8 %), einen Informatik- (1,8 %) oder einen technischen (0,1 %) Schulzweig an. Von sieben (0,6 %) TeilnehmerInnen wurde diesbezüglich keine Angaben gemacht und insgesamt gaben 68 (6,3 %) SchülerInnen einen anderen als die angeführten Schulzweige an.

Tabelle 9: Häufigkeitsverteilung und Anteilswerte der besuchten Schulzweige

Schulzweige		Häufigkeit	Prozent	Gültige Prozente
Gültig	Naturwissenschaftlich	247	22,6	22,9
	Sprachlich	288	26,3	26,7
	Darstellende Geometrie	46	4,2	4,3
	Bildnerisch	67	6,1	6,2
	Musikalisch	199	18,2	18,4
	Technisch	1	0,1	0,1
	Wirtschaftlich	27	2,5	2,5
	Humanistisch	20	1,8	1,9
	Informatik	20	1,8	1,8
	Sportlicher Zweig	68	6,2	6,3
	Sonstige	97	8,9	9,0
	Gesamt	1080	99,5	100,0
Fehlend	System	13	1,2	
Gesamt		1093	100,0	

3.3.2.4 Wiederholen einer Schulstufe

Im Anschluss gibt Tabelle 10 einen Überblick über die Häufigkeitsverteilungen im Wiederholen einer Schulstufe.

Tabelle 10: Häufigkeitstabelle Wiederholen einer Schulstufe

Wiederholen einer Schulstufe			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	921	543,5	377,5
ja	166	543,5	-377,5
Gesamt	1087		

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2 (1) = 524.402$, $p \leq .0001$ ein signifikantes Ergebnis. Das vorliegende Ergebnis weist auf einen Verteilungsunterschied im Wiederholen einer Schulstufe hin (siehe Tabelle 11).

Tabelle 11: Chi- Quadrat Test für Wiederholen einer Schulstufe

	Wiederholen
Chi-Quadrat	524,402 ^a
df	1
Asymptotische Signifikanz	$p \leq .0001$

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 543,5

Im Rahmen der Gesamtstichprobe (N = 1093) konnten insgesamt 1087 gültige Fälle ermittelt werden, von denen Angaben bezüglich dem Wiederholen einer Schulklasse vorliegen. Von 1087 SchülerInnen haben 921 (84,7 %) angegeben, noch nie eine Klasse wiederholt zu haben, hingegen berichteten 166 (15,3 %) SchülerInnen, schon zumindest einmal eine Klasse wiederholt zu haben. Von den insgesamt 403 Schülern gaben 77 (19,1 %) an eine Klasse wiederholt zu haben und von insgesamt 684 Schülerinnen konnten 89 (13 %) ermittelt werden, die ebenfalls zumindest einmal eine Klasse wiederholen mussten (siehe Tabelle 12).

Tabelle 12: Kontingenztafel Wiederholen einer Schulklasse in Abhängigkeit vom Geschlecht

Kreuztabelle- Geschlecht * Wiederholen					
Geschlecht			Wiederholen		Gesamt
			nein	ja	
	männlich	Anzahl	326	77	403
		Erwartete Anzahl	341,5	61,5	403,0
		% innerhalb von Geschlecht	80,9%	19,1%	100,0%
		Standardisierte Residuen	-,8	2,0	
	weiblich	Anzahl	595	89	684
		Erwartete Anzahl	579,5	104,5	684,0
		% innerhalb von Geschlecht	87,0%	13,0%	100,0%
		Standardisierte Residuen	,6	-1,5	
Gesamt		Anzahl	921	166	1087
		Erwartete Anzahl	921,0	166,0	1087,0
		% innerhalb von Geschlecht	84,7%	15,3%	100,0%

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2(1) = 7.28$, $p = .007$ ein signifikantes Ergebnis. Es können somit Verteilungsunterschiede in Bezug auf das Wiederholen einer Schulklasse in Abhängigkeit vom Geschlecht angenommen werden, männliche Schüler weisen mit 19,1 % einen höheren Anteil auf.

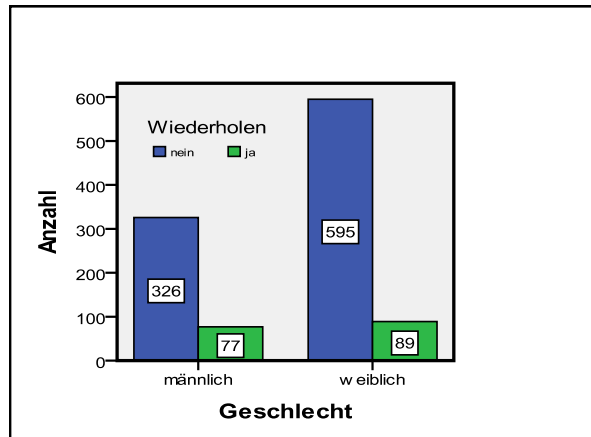


Abbildung 16. Verteilung Wiederholen in Abhängigkeit vom Geschlecht

In Abbildung 16 wird die Verteilung in Bezug auf das Wiederholen einer Schulklasse in Abhängigkeit vom Geschlecht graphisch dargestellt.

3.3.2.5 Wunsch nach einem Medizinstudium und Vorbereitung auf EMS

Folgende Tabelle gibt einen Überblick über die Häufigkeitsverteilung im Wunsch nach einem Medizinstudium (siehe Tabelle 13).

Tabelle 13: Häufigkeiten für den Wunsch auf ein Medizinstudium

Wunsch nach einem Medizinstudium			
	Beobachtetes N	Erwartete Anzahl	Residuum
nein	881	534,0	347,0
ja	187	534,0	-347,0
Gesamt	1068		

Die Berechnung mittel Chi- Quadrat Test $\chi^2 (1) = 450.970$, $p \leq .0001$ ein signifikantes Ergebnis, diesbezüglich können Verteilungsunterschiede im Wunsch auf ein Medizinstudium angenommen werden (siehe Tabelle 14).

Tabelle 14: Chi- Quadrat Test für den Wunsch auf ein Medizinstudium

	Medizin
Chi-Quadrat	450,970 ^a
df	1
Asymptotische Signifikanz	$p \leq .0001$

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 534,0.

Insgesamt konnten 1068 Personen ermittelt werden, von denen Angaben bezüglich des Wunsches, Medizin studieren zu wollen vorliegen. Diesbezüglich konnten 187 (17,5 %) SchülerInnen ermittelt werden, die die Frage nach dem Wunsch eines Medizinstudiums bejahten. Aus der folgenden Kreuztabelle kann im speziellen der Wunsch für ein Medizinstudium in Abhängigkeit vom Geschlecht entnommen werden. Insgesamt gaben 125 (18,6 %) weibliche Schüler und 62 (15,7 %) männliche Schüler an, nach dem Schulabschluss ein Medizinstudium beginnen zu wollen (siehe Abbildung 17).

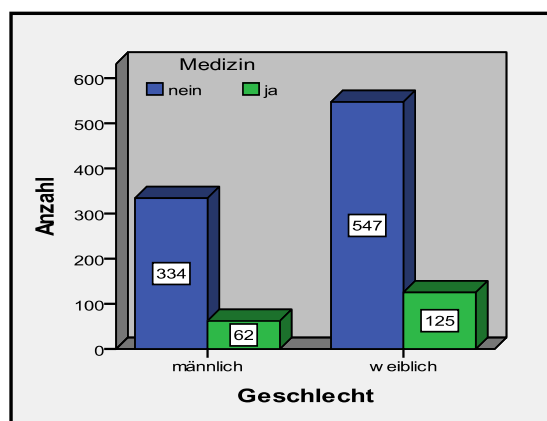


Abbildung 17 Wunsch auf ein Medizinstudium in Abhängigkeit vom Geschlecht

Aus der folgenden Kreuztabelle kann im speziellen der Wunsch auf ein Medizinstudium in Abhängigkeit vom Geschlecht entnommen werden. Insgesamt gaben 125 (18,6 %) weibliche Schüler und 62 (15,7 %) männliche Schüler an, nach dem Schulabschluss ein Medizinstudium beginnen zu wollen (siehe Tabelle 15).

Tabelle 15 Wunsch Medizin zu studieren in Abhängigkeit vom Geschlecht

Kreuztabelle - Geschlecht * Wunsch nach Medizinstudium				
Geschlecht		Medizin		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 entsprechenden Prüfgröße ergibt mit $\chi^2(1) = 1.496$, $p = .221$ ein nicht signifikantes Ergebnis. Somit kann kein Verteilungsunterschied im Wunsch, Medizin studieren zu wollen in Abhängigkeit vom Geschlecht beobachtet werden.

Auf Grundlage der folgenden Tabelle wird veranschaulicht, wie viele SchülerInnen, die ein Medizinstudium anstreben, sich auch schon mit dem EMS beschäftigt haben. Im Rahmen der Gesamtstichprobe konnten aus 1066 gültigen Fällen 117 (62,6 %) TeilnehmerInnen ermittelt werden, die den Wunsch auf ein Medizinstudium angaben und sich in irgendeiner Form mit dem EMS beschäftigt hatten. Die Zahl jener UntersuchungsteilnehmerInnen, die ein Medizinstudium beginnen möchten, sich aber noch nicht mit dem EMS beschäftigt hatten, umfasst 70 (37,4 %). Es ist zu beobachten, dass sich etwa zwei Drittel der SchülerInnen mit Wunsch auf ein Medizinstudium auch auf den EMS vorbereitet hat. Insgesamt 799 (90,9 %) der TeilnehmerInnen ohne Wunsch auf ein Medizinstudium hatten sich auch bisher noch nicht mit dem EMS auseinander gesetzt.

Tabelle 16 Vierfeldertafel Wunsch Medizin zu studieren und EMS- Beschäftigung

Kreuztabelle – Wunsch nach Medizinstudium * EMS - Beschäftigung					
			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 entsprechenden Prüfgröße $\chi^2(1) = 292.584$, $p \leq .0001$ zeigt ein signifikantes Ergebnis. Es kann ein Zusammenhang in der EMS-Beschäftigung und dem Wunsch, Medizin zu studieren angenommen werden.

Beschäftigung mit dem EMS und Geschlecht

Des Weiteren ist von Interesse, ob ein Zusammenhang in der Vorbereitung auf den EMS vom Geschlecht besteht. Von insgesamt 1085 gültigen Fällen gaben 142 (20,8 %) weibliche Testteilnehmer und 61 (15,2 %) männliche Testteilnehmer an, sich schon auf den EMS vorbereitet zu haben.

Insgesamt 882 (81,3 %) TeilnehmerInnen gaben an, sich nicht mit dem EMS beschäftigt zu haben, davon 541 (79,2 %) Schülerinnen und 341 (84,8 %) Schüler.

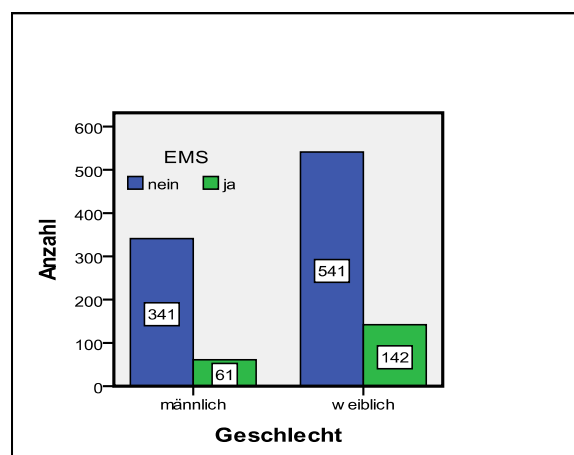


Abbildung 18: Beschäftigung mit dem EMS und Geschlecht

Aus der folgenden Tabelle wird die Vorbereitung auf den EMS in Abhängigkeit vom Geschlecht veranschaulicht.

Tabelle 17: Vierfeldertafel Beschäftigung mit EMS und Geschlecht

Kreuztabelle - Geschlecht * EMS – Beschäftigung					
Geschlecht			EMS		Gesamt
			nein	ja	
	männlich	Anzahl	341	61	402
		Erwartete Anzahl	326,8	75,2	402,0
		% innerhalb von Geschlecht	84,8%	15,2%	100,0%
		Standardisierte Residuen	,8	-1,6	
	weiblich	Anzahl	541	142	683
		Erwartete Anzahl	555,2	127,8	683,0
		% innerhalb von Geschlecht	79,2%	20,8%	100,0%
		Standardisierte Residuen	-,6	1,3	
Gesamt		Anzahl	882	203	1085
		Erwartete Anzahl	882,0	203,0	1085,0

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2(1) = 5.25$, $p = .002$ ein signifikantes Ergebnis. Aufgrund dieses Resultats kann ein signifikanter Verteilungsunterschied in der Vorbereitung mit dem EMS in Abhängigkeit vom Geschlecht angenommen werden. Es ist zu beobachten, dass sich Mädchen überdurchschnittlich häufiger mit einer Vorbereitung für den EMS beschäftigen.

Beschäftigung mit dem EMS und naturwissenschaftlicher Schulzweig

Ein weiterer bedeutender Aspekt ist, ob in Abhängigkeit vom schwerpunktmäßig gewählten naturwissenschaftlichen Schulzweig eine Beschäftigung mit dem EMS erfolgt. Von insgesamt 246 SchülerInnen, die einen naturwissenschaftlichen Schulzweig besuchten, hatten sich bereits 58 (23,6 %) mit einer Vorbereitung zum EMS beschäftigt. Von den 832 Personen, die keinen naturwissenschaftlichen Schulzweig besuchten,

hatten sich 144 (17,3 %) bereits mit dem EMS beschäftigt (siehe Abbildung 19).

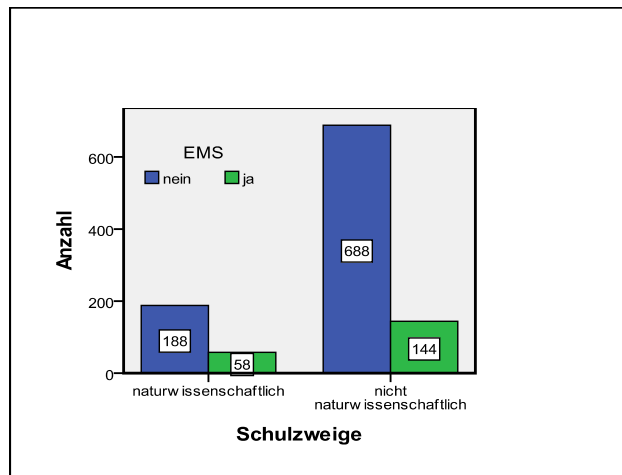


Abbildung 19: Beschäftigung mit EMS abhängig vom Schulzweig

Tabelle 18 gibt einen Überblick über die Vorbereitung zum EMS in Abhängigkeit des Schulzweiges. Von den 246 Personen, die einen naturwissenschaftlichen Schulzweig besuchen, geben 188 (76,4 %) TeilnehmerInnen an, sich nicht mit EMS- Vorbereitungen beschäftigt zu haben.

Tabelle 18: Vierfeldertafel Beschäftigung mit EMS und Schulzweig (naturwissenschaftlich)

Kreuztabelle – EMS Vorbereitung * Schulzweig				
Schulzweige		EMS		Gesamt
		nein	Ja	
naturwissenschaftlich	Anzahl	188	58	246
	Erwartete Anzahl	199,9	46,1	246,0
	% innerhalb von Schulzweig	76,4%	23,6%	100,0%
	Standardisierte Residuen	-,8	1,8	
nicht naturwissenschaftlich	Anzahl	688	144	832
	Erwartete Anzahl	676,1	155,9	832,0
	% innerhalb von Schulzweig	82,7%	17,3%	100,0%
	Standardisierte Residuen	,5	-1,0	
Gesamt	Anzahl	876	202	1078
	Erwartete Anzahl	876,0	202,0	1078,0
	% innerhalb von Schulzweig	81,3%	18,7%	100,0%

Die Berechnung der entsprechenden Prüfgröße zeigt mit $\chi^2(1) = 4.901$, $p = .027$ ein signifikantes Ergebnis. SchülerInnen mit naturwissenschaftlichem Schwerpunkt beschäftigen sich überzufällig häufiger mit dem EMS.

3.3.2.6 Verteilung der Geschlechter in die Versuchsbedingungen

In diesem Abschnitt wird die Verteilung der SchülerInnen in die Versuchsbedingungen, Stereotype Threat und Risikoverhalten beschrieben und veranschaulicht.

Stereotype Threat Bedingung in Abhängigkeit vom Geschlecht

Folgende Häufigkeitstabelle gibt einen Überblick der Zuteilung in die Stereotype Threat Bedingungen.

Tabelle 19: Häufigkeiten Stereotype Threat Bedingungen

Stereotype Threat			
	Beobachtetes N	Erwartete Anzahl	Residuum
A Aktivierung	293	273,3	19,8
B neutral	257	273,3	-16,3
C Aufklärung	280	273,3	6,8
D Negativaktivierung	263	273,3	-10,3
Gesamt	1093		

Die Berechnung der entsprechenden Prüfgröße $\chi^2(3) = 2.945$, $p = .400$ ein nicht signifikantes Ergebnis, somit können keine Verteilungsunterschiede in den Stereotype Threat Bedingungen angenommen werden.

Tabelle 20: Chi- Quadrat Test für Stereotype Threat Bedingungen

	Stereotype
Chi-Quadrat	2,945 ^a
df	3
Asymptotische Signifikanz	,400

- a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 273,3.

In Abbildung 20 werden die Häufigkeitsverteilungen in die vier Stereotype Threat - Bedingungen in Abhängigkeit vom Geschlecht dargestellt.

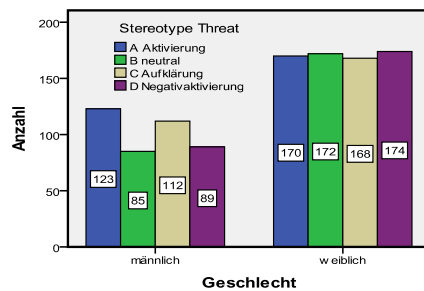


Abbildung 20: Verteilung der TeilnehmerInnen in die Stereotype Threat Bedingungen (Aktivierung, Neutral, Aufklärung und Negativaktivierung) in Abhängigkeit vom Geschlecht

Tabelle 21 zeigt die Zuordnung der SchülerInnen in die Stereotype Threat Versuchsbedingungen (Aktivierung, Neutral, Aufklärung und Negativaktivierung) in Abhängigkeit vom Geschlecht der TeilnehmerInnen.

Tabelle 21: Kontingenztafel Stereotype Threat und Geschlecht

Kreuztabelle – Stereotype Threat * Geschlecht						
Geschlecht		Stereotype Threat Bedingungen				Gesamt
		A	B	C	D	
männl.	Anzahl	123	85	112	89	409
	erwartet	109,6	96,2	104,8	98,4	409,0
	Stand. Residuen	1,3	-1,1	,7	-,9	
weibl.	Anzahl	170	172	168	174	684
	erwartet	183,4	160,8	175,2	164,6	684,0
	Stand. Residuen	-1,0	,9	-,5	,7	
gesamt	Anzahl	293	257	280	263	1093
	erwartet	293,0	257,0	280,0	263,0	1093,0

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2(3) = 6.909$, $p = .075$ ein nicht signifikantes Ergebnis. Die Zuweisung der UntersuchungsteilnehmerInnen in die Stereotype Threat-Bedingungen erfolgte unabhängig vom Geschlecht.

Risikoverhalten in Abhängigkeit vom Geschlecht

Nachfolgende Häufigkeitstabelle gibt einen Überblick über die Zuteilung der Versuchspersonen in die Risikobedingungen (siehe Tabelle 22).

Tabelle 22: Häufigkeiten Risikobedingungen

Risikoverhalten			
	Beobachtetes N	Erwartete Anzahl	Residuum
raten	373	364,3	8,7
ohne Instruktion	368	364,3	3,7
sicher	352	364,3	-12,3
Gesamt	1093		

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2(2) = .661$, $p = .719$ ein nicht signifikantes Ergebnis. Es können keine Verteilungsunterschiede in den Risikobedingungen angenommen werden (siehe Tabelle 23).

Tabelle 23: Chi- Quadrat Test für Risikobedingungen

	Risikoverhalten
Chi-Quadrat	,661 ^a
df	2
Asymptotische Signifikanz	,719

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 364,3.

Die nachfolgende Abbildung gibt einen Überblick über die Zuweisung der UntersuchungsteilnehmerInnen in die Risikobedingungen (Raten, Ohne Instruktion und Sicher) in Abhängigkeit vom Geschlecht. Wie auch in der

Stereotype Threat - Bedingung kann auch bei den Risikoverhaltens-Bedingungen eine Gleichverteilung angenommen werden.

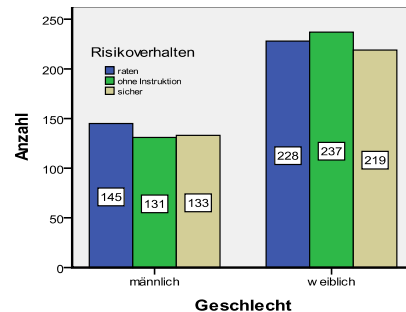


Abbildung 21: Zuweisung der TeilnehmerInnen in die Risikobedingungen (Raten, Ohne Instruktion und Sicher) in Abhängigkeit vom Geschlecht

Die nachfolgende Tabelle 24 veranschaulicht die Zuweisung der SchülerInnen in die vier Risikobedingungen (Raten, Ohne Instruktion und Sicher).

Tabelle 24: Kontingenztafel Risikoverhalten und Geschlecht

Kreuztabelle – Risikoverhalten * Geschlecht					
Geschlecht		Risikobedingungen			Gesamt
		raten	ohne Instruktion	sicher	
männlich	Anzahl	145	131	133	409
	Erwartete Anzahl	139,6	137,7	131,7	409,0
	Standardisierte Residuen	,5	-,6	,1	
weiblich	Anzahl	228	237	219	684
	Erwartete Anzahl	233,4	230,3	220,3	684,0
	Standardisierte Residuen	-,4	,4	-,1	
gesamt	Anzahl	373	368	352	1093
	Erwartete Anzahl	373,0	368,0	352,0	1093,0

Die Berechnung der entsprechenden Prüfgröße weist mit $\chi^2(2) = .878$, $p = .645$ auf ein nicht signifikantes Ergebnis hin. Die Zuteilung der SchülerInnen in die Risikobedingungen (*Sicher, Ohne Instruktion und Raten*) erfolgte somit unabhängig vom Geschlecht.

Stereotype Threat- Bedingungen und Risikoinstruktion

Folgende Abbildung gibt einen Überblick der Verteilungen der SchülerInnen in die Risikobedingung in Abhängigkeit von der Stereotype Threat-Bedingung.

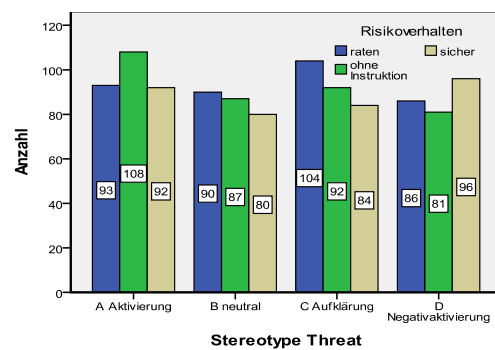


Abbildung 22: Stereotype Threat und Risikoinstruktion

Die anschließende Tabelle zeigt die Verteilungen der SchülerInnen in den drei Risikobedingungen in Abhängigkeit der vier Stereotype Threat-Bedingungen (siehe Tabelle 25).

Tabelle 25: Kontingenztafel Stereotype Threat und Risikoverhalten

Stereotype * Risikoverhalten Kreuztabelle					
Stereotype Threat		Risikoverhalten			Gesamt
		raten	ohne Instruktion	Sicher	
A	Anzahl	93	108	92	293
	Erwartete Anzahl	100,0	98,6	94,4	293,0
	Standardisierte Residuen	-,7	,9	-,2	
B	Anzahl	90	87	80	257
	Erwartete Anzahl	87,7	86,5	82,8	257,0
	Standardisierte Residuen	,2	,1	-,3	
C	Anzahl	104	92	84	280
	Erwartete Anzahl	95,6	94,3	90,2	280,0
	Standardisierte Residuen	,9	-,2	-,7	
D	Anzahl	86	81	96	263
	Erwartete Anzahl	89,8	88,5	84,7	263,0
	Standardisierte Residuen	-,4	-,8	1,2	
Gesamt	Anzahl	373	368	352	1093
	Erwartete Anzahl	373,0	368,0	352,0	1093,0

Die Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 5.121$, $p = .528$ ergibt ein nicht signifikantes Ergebnis. Unterschiede in Bezug auf die Verteilung im Risikoverhalten in Abhängigkeit von Stereotype Threat können ausgeschlossen werden. Die Zuteilung der TeilnehmerInnen erfolgte unabhängig von den Stufen der experimentellen Faktoren.

3.3.2.7 Deskriptivstatistik der EMS – ähnlichen Aufgaben

Im Folgenden wird auf die Verteilungen der einzelnen EMS-Bereiche eingegangen.

Eine Normalverteilung der Daten in den sechs EMS-ähnlichen Untertests über die Gesamtstichprobe (N = 1093) kann angenommen werden. Desweiteren kann man der Tabelle 26 die Reliabilitäten der EMS-ähnlichen Untertests entnehmen.

Tabelle 26: Statistische Kennwerte und Koeffizienten der EMS- ähnlichen Untertests

Statistik der EMS- ähnlichen Untertests						
N = 1093	EMS Quant	EMS Schlauch	EMS Natur	EMS Diagramme	EMS Figuren	EMS Fakten
Mittelwert	3.00	10.19	2.81	3.01	4.43	4.24
Median	3.00	10.00	3.00	3.00	4.00	4.00
Itemanzahl	16	21	16	16	8	8
Stand.abweichung	1.675	3.446	1.688	1.633	2.053	1.767
Schiefe	.485	-.168	.482	.370	-.068	-.111
Kurtosis	.155	-.283	-.150	-.037	-.638	-.426
Minimum	0	0	0	0	0	0
Maximum	10	20	9	9	8	8
Cronbach's Alpha	.335	.651	.353	.314	.637	.489

3.3.2.8 Relative Lösungshäufigkeiten in Abhängigkeit der Stichprobe

Im Folgenden werden die Lösungshäufigkeiten und die entsprechenden Standardabweichungen der beiden Untertests *Fakten Lernen* und *Figuren Lernen* dargestellt. Die Ratewahrscheinlichkeit liegt für die TestteilnehmerInnen bei einem Fünftel, gleichbedeutend mit einer Wahrscheinlichkeit von 20 %, die Lösung einer Aufgabe zu erraten. Diese wird bei der Berechnung der Itemschwierigkeit nicht berücksichtigt.

Tabelle 27: Lösungswahrscheinlichkeiten im Untertest Figuren Lernen

UT-Figuren Lernen	<i>M</i> Gesamt	<i>SD</i> Gesamt	<i>M</i> Männer	<i>SD</i> Männer	<i>M</i> Frauen	<i>SD</i> Frauen
ReprodFig47	0.59	0.491	0.57	0.496	0.61	0.489
ReprodFig48	0.57	0.495	0.57	0.495	0.58	0.495
ReprodFig49	0.70	0.457	0.70	0.460	0.71	0.456
ReprodFig50	0.43	0.495	0.39	0.489	0.45	0.498
ReprodFig51	0.38	0.486	0.38	0.486	0.38	0.485
ReprodFig52	0.70	0.460	0.69	0.461	0.70	0.460
ReprodFig53	0.42	0.494	0.43	0.495	0.42	0.494
ReprodFig54	0.63	0.482	0.63	0.484	0.64	0.481

Die relative Lösungshäufigkeit liegt für den Untertest *Figuren Lernen* im mittleren Bereich.

Tabelle 28: Lösungswahrscheinlichkeiten im Untertest Fakten Lernen

UT-Fakten Lernen	<i>M</i> Gesamt	<i>SD</i> Gesamt	<i>M</i> Männer	<i>SD</i> Männer	<i>M</i> Frauen	<i>SD</i> Frauen
ReprodFakt55	0.65	0.477	0.63	0.483	0.66	0.474
ReprodFakt56	0.43	0.495	0.37	0.484	0.46	0.499
ReprodFakt57	0.35	0.478	0.35	0.479	0.35	0.478
ReprodFakt58	0.52	0.500	0.47	0.500	0.54	0.498
ReprodFakt59	0.85	0.355	0.84	0.368	0.86	0.348
ReprodFakt60	0.39	0.489	0.35	0.479	0.42	0.493
ReprodFakt61	0.65	0.477	0.57	0.495	0.69	0.461
ReprodFakt62	0.40	0.490	0.34	0.474	0.43	0.496

3.4 AUSWERTUNG UND ERGEBNISSE

Dieses Kapitel beschäftigt sich mit der testtheoretischen Überprüfung der Wirksamkeit der experimentellen Variationen. Es wird die Wirksamkeit der instruierten vier Stereotype Threat-Bedingungen überprüft, welche in

Anlehnung an dem in der Studie von Halpern und Tan (2010) vorgegeben Wissenstest erhoben wurden.

Im Anschluss daran erfolgt eine Darstellung der Reliabilitätsanalysen der EMS-ähnlichen Untertests und des Game of Dice Task (GDT, Brand et al., 2005).

Daran anschließend werden die Ergebnisse der Haupthypothesen, als auch der interessierenden Nebenhypothesen im Detail dargestellt.

Zu Beginn wird ein Überblick über die in der statistischen Datenanalyse eingesetzten Verfahren gegeben. Im Anschluss daran soll auf diese nicht mehr detailliert eingegangen werden, es soll nur mehr darauf hingewiesen werden, welches Auswertungsverfahren für die spezifische Fragestellung verwendet wurde.

3.4.1 AUSWERTUNGSVERFAHREN

Im Rahmen der statistischen Analyse der Daten kommen folgende Verfahren zum Einsatz (vgl. Backhaus, Erichson, Plinke & Weiber, 2008; Bortz & Döring, 2006):

❖ Varianzanalyse (ANOVA)

Dieses allgemeine lineare Modell prüft den Einfluss von einer oder mehreren nominalskalierten unabhängigen Variablen auf eine intervallskalierte abhängige Variable.

❖ Levene- Test

Mittels Levene-Test wird die Gleichheit der Varianzen in unterschiedlichen Stichproben untersucht. Der Levene-Test muss ein nicht signifikantes Ergebnis zeigen, es kann somit die Nullhypothese beibehalten werden, um die Homogenität der Varianzen annehmen zu können.

❖ **Kontingenztafeln bzw. Kreuztabellen und χ^2 - Test**

Eine Kontingenztafel untersucht mögliche Zusammenhänge zwischen zwei nominalskalierten Merkmalen. Mittels Kontingenztafeln wird untersucht ob sich die empirischen von den theoretischen Häufigkeiten unterscheiden. Desweiteren kann mittels χ^2 – Test untersucht werden, ob sich die beobachteten Häufigkeiten signifikant von den zu erwarteten Häufigkeiten unterscheiden.

❖ **Produkt- Moment- Korrelation nach Pearson**

Diese Korrelation untersucht die Stärke und die Richtung von zwei intervallskalierten Merkmalen unter der Bedingung eines linearen Zusammenhangs. Eine weitere Voraussetzung ist die Normalverteilung der Daten.

❖ **A- priori Kontraste**

Die Voraussetzung bei A-priori Kontrasten ist das Formulieren von Einzelvergleichshypothesen um mögliche signifikante Unterschiede zu identifizieren.

3.4.2 ÜBERPRÜFUNG DER EXPERIMENTELLEN VARIATIONEN MITTELS WISSENSTEST

Nachfolgend wird die testtheoretischen Überprüfung und die resultierenden Ergebnisse der Wirksamkeit der experimentellen Versuchsbedingung Stereotype Threat dargestellt.

Überprüfung der experimentellen Stereotype Threat- Bedingungen

Im Anschluss an den soziodemographischen Fragebogenteil und noch vor der Darbietung der sechs EMS-ähnlichen Untertests wurde den VersuchsteilnehmerInnen der Wissenstest vorgelegt, mit welchem die experimentellen Variationen der vier Stereotype Threat-Bedingungen erfasst werden sollen. Der Wissenstest wurde der Studie von Halpern und Tan (2001), *Beliefs About the Cognitive Abilities of Females and Males*

Questionnaire entnommen. Liegen die Prozentmittelwerte innerhalb des Bereichs von 45,0 und 55,0, so entsprechen sie den Stereotype Threat Bedingungen *Neutral* und *Aufklärung*; sind Mittelwerte beobachtbar, die außerhalb dieses Bereichs liegen, kann eine Wirksamkeit der Positivaktivierung und der Negativaktivierung der Stereotype Threat-Bedingungen angenommen werden.

Zum besseren Verständnis sollen an dieser Stelle die vier Stereotype Threat-Instruktionen kurz dargestellt werden:

- *Aktivierung von Stereotype Threat*

„Frauen zeigen bessere Leistungen in Gedächtnisaufgaben, Männer hingegen in kognitiven Aufgaben die das Arbeiten mit Zahlen, die räumliche Orientierung, das Verstehen von komplexen Texten und das naturwissenschaftliche Verständnis betreffen“.

- *Neutrale Stereotype Threat Bedingung*

„Es gibt keine geschlechtsspezifischen Unterschiede in der Bearbeitung von kognitiven Aufgaben“.

- *Aufklärung Stereotype Threat Bedingung*

„Die Aufklärung von Stereotype Threat, als auch die möglichen Ursachen auf die Testsituation wurden erklärt“.

- *Negativaktivierung von Stereotype Threat*

„Frauen zeigen bessere Leistungen im Bearbeiten von Aufgaben, die räumliche Orientierung, das Verstehen von komplexen Texten, Arbeiten mit Zahlen und naturwissenschaftliches Verständnis betreffen und Männer bessere Leistungen in Gedächtnisaufgaben“.

Im Folgenden wird die Wirksamkeit der vier Stereotype Threat Bedingungen mittels χ^2 -Tests für die einzelnen EMS-ähnlichen Aufgaben berechnet.

Im Detail seien hier nur die für die vorliegende Arbeit interessierenden Untertests *Fakten Lernen* und *Figuren Lernen* näher beschrieben. Im

Anhang sind die entsprechenden Kontingenztafeln der weiteren vier EMS-ähnlichen Untertests zu finden.

Ergänzend zu den sechs EMS-ähnlichen Aufgaben wurde sowohl zu den Fremdeinschätzungen, als auch zu den Selbsteinschätzungen des Wissenstests eine Aufgabe zu Theory of Mind vorgelegt, um diesbezüglich auch Informationen über die Wirksamkeit von Stereotype Threat zu erlangen.

Tabelle 29 zeigt die Stereotype Threat-Bedingungen (Aktivierung, Neutral, Aufklärung und Negativaktivierung) in Abhängigkeit der Wissenstest-Aufgabe Fakten Lernen.

Tabelle 29: Kontingenztafel Stereotype Threat (4) und Wissenstest Aufgabe Fakten Lernen

Kreuztabelle Stereotype * Fakten Lernen						
			Fakten			Gesamt
			nein	ja	entgegen	
Stereotype	Aktivierung	Anzahl	48	205	37	290
		erwartet	82,3	155,5	52,3	290,0
		% innerhalb von	16,6%	70,7%	12,8%	100,0%
		Stand. Residuen	-3,8	4,0	-2,1	
	Neutral	Anzahl	84	132	38	254
		erwartet	72,1	136,2	45,8	254,0
		% innerhalb von	33,1%	52,0%	15,0%	100,0%
		Stand. Residuen	1,4	-,4	-1,1	
	Aufklärung	Anzahl	127	104	47	278
		erwartet	78,9	149,0	50,1	278,0
		% innerhalb von	45,7%	37,4%	16,9%	100,0%
		Stand. Residuen	5,4	-3,7	-,4	
	Negativaktivierung	Anzahl	48	139	73	260
		erwartet	73,8	139,4	46,9	260,0
		% innerhalb von	18,5%	53,5%	28,1%	100,0%
		Stand. Residuen	-3,0	,0	3,8	
Gesamt		Anzahl	307	580	195	1082
		erwartet	307,0	580,0	195,0	1082,0
		% innerhalb von	28,4%	53,6%	18,0%	100,0%

Aufgrund der Berechnung der entsprechenden Prüfgröße $\chi^2(6) = 104.697$, $p \leq .0001$ kann ein signifikantes Ergebnis beobachtet werden. Es können signifikante Verteilungsunterschiede im Wissenstest-Fremdeinschätzungsaufgabe Fakten Lernen in Abhängigkeit von den Stereotype Threat-Bedingungen angenommen werden.

Tabelle 30: Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe Figuren Lernen

Kreuztabelle Stereotype * Figuren Lernen						
			Figuren			Gesamt
			nein	ja	entgegen	
Stereotype	Aktivierung	Anzahl	115	124	50	289
		erwartet	76,8	172,2	40,0	289,0
		% innerhalb von	39,8%	42,9%	17,3%	100,0%
		Stand. Residuen	4,4	-3,7	1,6	
	Neutral	Anzahl	37	205	13	255
		erwartet	67,7	152,0	35,3	255,0
		% innerhalb von	14,5%	80,4%	5,1%	100,0%
		Stand. Residuen	-3,7	4,3	-3,8	
	Aufklärung	Anzahl	33	216	30	279
		erwartet	74,1	166,3	38,6	279,0
		% innerhalb von	11,8%	77,4%	10,8%	100,0%
		Stand. Residuen	-4,8	3,9	-1,4	
	Negativaktivierung	Anzahl	103	101	57	261
		erwartet	69,3	155,5	36,1	261,0
		% innerhalb von	39,5%	38,7%	21,8%	100,0%
		Stand. Residuen	4,0	-4,4	3,5	
Gesamt		Anzahl	288	646	150	1084
		erwartet	288,0	646,0	150,0	1084,0
		% innerhalb von	26,6%	59,6%	13,8%	100,0%

Die Berechnung der entsprechenden Prüfgröße $\chi^2 (6) = 168.720$, $p \leq .0001$ ergibt ein signifikantes Ergebnis. Somit können Verteilungsunterschiede in der Wissenstest-Fremdeinschätzungsaufgabe *Figuren Lernen* in Abhängigkeit von den Stereotype Threat-Bedingungen festgestellt werden.

Aufgrund der Ergebnisse des Wissenstests mittels χ^2 -Tests kann die Überprüfung der Wirksamkeit der Stereotype Threat-Bedingungen für alle Wissenstest-Aufgaben angenommen werden.

Des Weiteren können auch die Prozentmittelwerte zur Überprüfung der Wirksamkeit der Stereotype Threat Bedingungen herangezogen werden (siehe Tabelle Anhang).

Je nach Fähigkeitsbereich ist erkennbar, dass die Stereotype Threat-Bedingungen an sich wirksam sind, jedoch in Abhängigkeit der abhängigen Variablen und unter Berücksichtigung der standardisierten Residuen unterschiedliche Richtungen annehmen können. Sobald ein standardisiertes Residuum einen Wert im Bereich $\geq \pm 2$ aufweist, kann ein Unterschied zwischen beobachteter und erwarteter Häufigkeit angenommen werden (Bühl & Zöfel, 2000).

Als zweiter Teil des Wissenstests wurden sieben Selbsteinschätzungsaufgaben vorgegeben, welche mit einer zweifaktoriellen multivariaten Varianzanalyse in Abhängigkeit vom Geschlecht und der vier Stereotype Threat - Bedingungen geprüft werden. Obwohl die Homogenität der beiden Selbsteinschätzungsaufgaben *Figuren Lernen* und *Schlauchfiguren* ($p < .05$) nicht anzunehmen ist, erfolgt die Hypothesenprüfung aufgrund der Robustheit des Verfahrens mittels varianzanalytischem Design.

In die Analysen können 1065 gültige Fälle miteinbezogen werden. In Bezug auf die Unterschiede im Geschlecht und abhängig von den Stereotype Threat-Bedingungen werden die Ergebnisse durch Tabelle 31 veranschaulicht.

Tabelle 31: Kennwerte und Prüfgrößen der Selbsteinschätzung im Wissenstest in Abhängigkeit vom Geschlecht

EMS- ähnliche Untertests	Geschlecht	<i>M</i>	<i>SD</i>	<i>F</i> (1,1057)	Sig.
Figuren Lernen	männlich	6.63	0.81	4.677	.031*
	weiblich	6.50	1.04		
Fakten Lernen	männlich	4.93	1.30	1.753	.186
	weiblich	5.05	1.23		
Med-naturwiss. Verständnis	männlich	5.09	1.43	5.277	.022*
	weiblich	4.86	1.39		
Diagramme und Tabellen	männlich	5.09	1.57	73.546	≤ .0001**
	weiblich	4.22	1.60		
Quantitative und Formale Probleme	männlich	5.05	1.68	78.300	≤ .0001**
	weiblich	4.04	1.79		
Schlauchfiguren	männlich	5.89	1.40	48.401	≤ .0001**
	weiblich	5.16	1.73		

Männliche Schüler schätzen ihre Leistungen in fünf EMS-ähnlichen Bereichen im Rahmen der Selbsteinschätzung des Wissenstests, *Diagramme und Tabellen* $F(1, 1057) = 73.546, p \leq .0001$; *Quantitative und Formale Probleme* $F(1,1057) = 78.300, p \leq .0001$; *Schlauchfiguren* $F(1,1057) = 48.401, p \leq .0001$; *Figuren Lernen* $F(1,1057) = 4.677, p = .031$; *Medizinisch-naturwissenschaftliches Grundverständnis* $F(1,1057) = 5.277, p = .022$ signifikant höher ein als Mädchen.

Im Bereich *Fakten Lernen* $F(1,1057) = 1.753, p = .186$ können keine signifikanten Unterschiede in den Einschätzungen in Bezug auf das Geschlecht festgestellt werden.

Signifikante Interaktionen in den Selbststeinschätzungen des Wissenstest unter Einbeziehung von Geschlecht und Stereotype Threat-Bedingungen

sind in den Aufgaben *Diagramme und Tabellen* $F(1,1057) = 6.820$, $p \leq .0001$ festzustellen.

In Abbildung 23 ist die Interaktion im Wissenstest Selbsteinschätzungsaufgaben *Diagramme und Tabellen* dargestellt. Unterschiede zwischen weiblichen und männlichen Schülern in der Einschätzung in Bezug auf ihre Leistungen können in allen Stereotype-Threat-Bedingungen beobachtet werden.

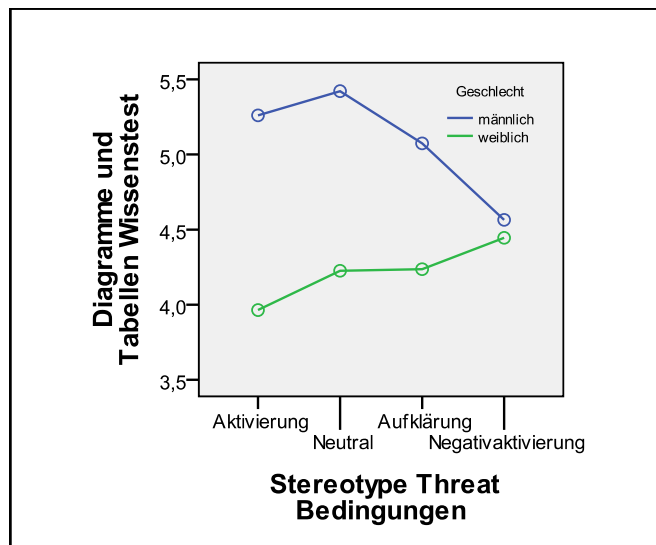


Abbildung 23: Wissenstest Aufgabe Diagramme und Tabellen

Abschließend kann festgehalten werden, dass die Überprüfung der Wirksamkeit der Stereotype Threat Bedingungen (Aktivierung, Neutral, Aufklärung und Negativaktivierung), welche durch den Wissenstest erfolgte, als wirksam angenommen werden kann.

Desweiteren kann bei den Selbsteinschätzungen festgestellt werden, dass die Stereotype Threat - Bedingungen im Untertest *Diagramme und Tabellen* Auswirkungen bewirkt hat.

3.4.3 RELIABILITÄTSANALYSEN

Im Folgenden werden die Reliabilitäten der Skalen der EMS-ähnlichen Untertests und des computerisierten Verfahrens *Game of Dice Task* (GDT) behandelt. Darüber hinaus wird eine Reliabilitätsanalyse mittels Cronbach's Alpha für den EMS-ähnlichen Gesamttest als auch eine Reliabilitätsanalyse des EMS-ähnlichen Gesamttest für die Dimension Merkfähigkeit durchgeführt.

3.4.3.1 Reliabilitätsanalysen der EMS-ähnlichen Aufgaben und Game of Dice Task (GDT) mittels Cronbach's Alpha

Die *Reliabilität* eines Tests erklärt den Grad der Genauigkeit, mit dem ein Test ein Merkmal erfasst bzw. misst. Die Angabe der inneren Konsistenz kann unter anderem mittels Cronbach's Alpha erfolgen, das Ergebnis ist eine Abschätzung der unteren Grenze der Reliabilität (Kubinger, 2006).

Tabelle 32: Vergleich Cronbach's Alpha Reliabilitäten der EMS Untertests (Hänsgen & Spicher, 2007) und Game of Dice Task (GDT)

EMS-ähnliche Untertests (N = 1093)	Cronbach's Alpha	Cronbach's Alpha (Hänsgen & Spicher, 2007)
Figuren Lernen	.637	.63
Fakten Lernen	.489	.73
Quantitative u. formale Probleme	.335	.70
Diagramme und Tabellen	.314	.69
Med.- naturwiss. Grundverständnis	.353	.65
Schlauchfiguren	.651	.73
Game of Dice Task (GDT, N = 471)	.891	-

Die Reliabilitäten für die dichotomen Items der sechs EMS- ähnlichen Untertests werden mittels Cronbach-Alpha-Koeffizienten ermittelt.

Cronbach's Alpha zur Beurteilung der Reliabilität hat den Vorteil, dass die Items auch unterschiedliche Schwierigkeiten aufweisen können. Da die Items unterschiedliche Schwierigkeitsgrade aufweisen, ist die Methode nach Cronbach's Alpha, die nur essentielle *tau*-Äquivalenz voraussetzt, die Methode nach Wahl (Rost, 2004, S. 380).

Die Resultate der Analyse der Reliabilitäten mittels Cronbach's Alpha machen deutlich, dass von den insgesamt sechs EMS-ähnlichen Untertests, welche im Rahmen dieser Untersuchung vorgegeben wurden, nur zwei Untertests, *Figuren Lernen* (.637) und *Schlauchfiguren* (.651) über eine ausreichend hohe Reliabilität verfügen.

Die restlichen Untertest der EMS- ähnlichen Aufgaben weisen geringe und unbefriedigende Reliabilitäten auf und können als wenig konsistent bezeichnet werden. Wenn eine Skala ein niedriges Cronbach's Alpha aufweist, ist dies ein Hinweis dafür, dass mittels dieser Untertests nicht nur eine einzige Dimension bzw. Fähigkeit erfasst wird und gleichzeitig die Skalierung als nicht zufriedenstellend angenommen werden kann.

Das Nebengütekriterium *Skalierung* ist dann erfüllt, wenn die resultierenden Werte eines Tests die empirischen Verhaltensrelationen adäquat abbilden (Kubinger, 2006, S. 79). Die Zielsetzung ist, möglichst eindimensional zu messen, womit die innere Konsistenz eines Tests hoch ist. Die einzelnen Items eines Tests sollen somit gemeinsam auf einen Faktor laden. Die *Speed- and- Power Testung* stellt hier ein klassisches Problem dar, wodurch zwei Fähigkeiten vermisch werden; einerseits sollen bestimmte Aufgaben erfüllt bzw. gelöst werden und andererseits soll auch die Fähigkeit vorhanden sein, dies möglichst bzw. ausreichend schnell zu tun (Kubinger, 2006).

Die zu beobachtenden Unterschiede bei der Bearbeitung der Untertests in Abhängigkeit vom Geschlechter machen deutlich, dass männliche Schüler zwar z.B. im Untertest *Quantitative und Formale Probleme* besser abschneiden, die nicht hinreichend vorhandene Skalierung der Skalen gibt jedoch Anlass zu vermuten, dass nicht nur die intendierte Fähigkeit

gemessen wird. Dies lässt vermuten, dass männliche Schüler eventuell unter anderem deswegen besser abschneiden, weil sie mit der Speed-Komponente im Test besser umgehen können als Mädchen.

Die Reliabilitätsanalyse mittels Cronbach's Alpha zum *Game of Dice Task* (GDT) ergibt eine ausreichend hohe und zufriedenstellende Reliabilität.

3.4.3.2 Reliabilitätsanalysen der Faktorstufenkombinationen

Stereotype Threat und Risikoverhalten der 6 EMS- ähnlichen Untertests

Folgende Tabellen (33-38) geben einen Überblick über die Reliabilitäten der Faktorstufenkombinationen Stereotype Threat und Risikoverhalten der sechs EMS- ähnlichen Untertests.

Tabelle 33: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Untertest Figuren Lernen

Risiko/Stereo	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; 8)

Es zeigt sich dass die Faktorstufenkombination Stereotype Threat *Negativaktivierung* (D) und Risikoinstruktion *Sicher* im Untertest Figuren Lernen die höchste Reliabilität mit .752 aufweist (siehe Tabelle 33).

Tabelle 34: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) 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)

Aus Tabelle 21 ist ersichtlich, dass die Faktorstufenkombination *Stereotype Threat Aktivierung (A)* und Risikoinstruktion *Ohne Instruktion* zur höchsten Reliabilität mit .607 führt.

Tabelle 35: 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)

Die Faktorstufenkombination *Stereotype Threat Neutral* und Risikoinstruktion *Sicher* bedingen die höchste Reliabilität (.432) im Untertest Quantitative und Formale Probleme (siehe Tabelle 35).

Tabelle 36: 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)

Aus Tabelle 36 ist ersichtlich, dass die Kombination der Stereotype Threat Bedingung *Neutral* und die Risikoinstruktionsbedingung *Ohne Instruktion* im Untertest Medizinisch-naturwissenschaftliches Grundverständnis die höchste Reliabilität mit .476 bedingt.

Tabelle 37: 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)

Die Faktorstufenkombination der Stereotype Threat Bedingung *Aufklärung* und der Risikoinstruktion *Sicher* ergeben für den Unterst Diagramme und Tabellen mit .434 die höchste Reliabilität (siehe Tabelle 37).

Aus der folgenden Tabelle 38 ist zu beobachten, dass im Untertest Schlauchfiguren die Stereotype Threat Bedingung *Aktivierung* und die Risikoinstruktionsbedingung *Raten* zur höchsten Reliabilität .728 führt.

Tabelle 38: Reliabilitäten der Faktorstufenkombination (Stichprobenumfang, Itemanzahl) EMS- ähnlicher Unterst 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)

3.4.3.3 Reliabilitätsanalyse des EMS- ähnlichen Gesamttest und der Untertests zur Erfassung der Merkfähigkeit

Da aus testtheoretischer Sicht die EMS- ähnlichen Untertests als beliebige Testteile aufgefasst werden können, wird diesbezüglich der Frage nachgegangen, ob für den EMS-ähnlichen Gesamttest, auch eine hohe Gesamt-Reliabilität anzunehmen ist.

Tabelle 39: Reliabilität mittels Cronbach's Alpha EMS- ähnlicher Gesamttest

Cronbach's Alpha	N
.532	1093

Aus der Tabelle 20 ist ersichtlich, dass die Gesamt- Reliabilität mittels Cronbach's Alpha = .532 relativ niedrig ausfällt. Es kann angenommen werden, dass mit den sechs EMS- ähnlichen Untertests unterschiedliche

Dimensionen erfasst werden und keine Hinweise bestehen, einen EMS-Gesamtscore zu bilden.

Im Folgenden wird eine Reliabilitätsanalyse zur Berechnung des Cronbach's Alpha für die EMS- ähnlichen Untertest (Figuren- und Fakten Lernen) zur Erfassung der Merkfähigkeit durchgeführt, um zu prüfen, ob die Untertests als ein Gesamtscore zusammengefasst werden können (siehe Tabelle 40).

Tabelle 40: Reliabilitäten mittels Cronbach's Alpha EMS- ähnlicher Gesamttest

Cronbach's Alpha	N
.428	1093

Aufgrund des Cronbach's Alpha-Koeffizienten von .428 besteht kein Grund, die beiden EMS- ähnlichen Untertests als Gesamtscore zusammenzufassen. Die beiden Untertest erfassen offenbar unterschiedliche Fähigkeiten.

3.4.4 INTERKORRELATIONEN DER EMS- ÄHNLICHEN UNTERTESTS

Im Folgenden werden die Zusammenhänge der einzelnen EMS- ähnlichen Untertests mittels einer Produkt- Moment- Korrelation nach Pearson dargestellt.

Insgesamt können 15 mögliche paarweise Interkorrelationen betrachtet werden.

Tabelle 41: Interkorrelationsmatrix der Produkt- Moment Korrelation r der EMS- ähnlichen Untertests (N = 1093)

	Quant	Schlauch	Natur	Diagramme	Figuren
EMS Schlauch	.162**				
EMS Natur	.273**	.224**			
EMS Diagramme	.287**	.170**	.334**		

EMS Figuren	.067*	.178**	.077*	.094**	
EMS Fakten	.110**	.169**	.180**	.116**	.276**

Es ist erkennbar, dass alle EMS-ähnlichen Untertests einen positiven Zusammenhang aufweisen. Die Korrelationskoeffizienten weisen auf geringe bis mittelhohe Zusammenhänge hin. Der höchste Zusammenhang kann zwischen den Fakten Lernen und Figuren Lernen ist mit $r = .276$ beobachtet werden. Es kann davon ausgegangen werden, dass mit zunehmender Fähigkeit der Person auch eine entsprechend gute Leistung in den meisten EMS-ähnlichen Untertests zu erwarten ist.

3.4.5 DARSTELLUNG DER HYPOTHESENBEZOGENEN ERGEBNISSE

Im folgenden Teil soll die statistische Hypothesenprüfung, der Haupt- und Nebenfragestellungen zu den EMS-ähnlichen Untertests *Fakten Lernen* und *Figuren Lernen* dargestellt werden.

Vor der Überprüfung der statistischen Hypothesen soll eine Auflistung der jeweils interessierenden Hypothesen im Rahmen eines Hypothesenblocks verdeutlicht werden.

Zur Prüfung der Haupt- und Nebenhypothesen wird eine univariate zwei- bzw. dreifaktorielle Varianzanalyse (ANOVA) berechnet. Die Auswertungsverfahren, welche im Rahmen dieser Analyse zum Einsatz kamen, wurden bereits in einem vorherigen Kapitel ausführlich beschrieben. In diesem Abschnitt wird nur mehr daraufhin gewiesen, welches Auswertungsverfahren für die entsprechende Fragestellung verwendet wurde.

In einem weiteren Teil dieses Kapitels sollen die Nebenhypothesen des *Game of Dice Task* (GDT) beschrieben werden.

3.4.5.1 Darstellung der univariaten Ergebnisse des EMS- ähnlichen Untertests: Figuren Lernen und Fakten Lernen

Hypothesenblock I.

H₁(1) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Geschlecht

H₁(2) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit der Stereotype Threat- Bedingungen (4)

H₁(3) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Risikoverhalten (3)

H₁(4) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Geschlecht und der Stereotype Threat- Bedingungen(4)

H₁(5) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Geschlecht und dem Risikoverhalten (3)

H₁(6) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit der Stereotype Threat- Bedingungen (4) und dem Risikoverhalten (3)

H₁(7) Es gibt eine signifikante Interaktion höherer Ordnung im EMS- ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Geschlecht, der Stereotype Threat- Bedingungen (4) und dem Risikoverhalten (3)

Die Hypothesenprüfung erfolgte mit einer univariaten dreifaktoriellen Varianzanalyse (ANOVA).

In der nachfolgenden Hypothesenprüfung werden drei unabhängige Variablen bzw. Faktoren, nämlich Geschlecht, Stereotype Threat Bedingungen und Risikoverhalten bezüglich einer abhängigen Variable, dem EMS- ähnlichen Untertest *Figuren Lernen*, berücksichtigt (siehe Tabelle 42).

Tabelle 42: Kennwerte EMS- ähnlicher Untertest Figuren Lernen

		männlich			Weiblich			Gesamt		
Stereo	Risiko	<i>M</i>	<i>SD</i>	n	<i>M</i>	<i>SD</i>	n	<i>M</i>	<i>SD</i>	n
A	Gesamt	4.4	1.974	12	4.3	2.01	17	4.38	1.992	293
B	Gesamt	4.0	2.098	85	4.3	2.07	17	4.26	2.083	257
C	Gesamt	4.4	2.026	11	4.6	1.93	16	4.56	1.972	280
D	Gesamt	4.5	2.292	89	4.5	2.10	17	4.51	2.16	263
Gesamt	Raten	4.5	1.935	14	4.3	1.99	22	4.4	1.96	373
	Ohne	4.0	2.138	13	4.8	2.02	23	4.5	2.09	368
	Sicher	4.4	2.171	13	4.2	2.05	21	4.3	2.09	352
	Gesamt	4.3	2.085	40	4.4	2.03	68	4.43	2.05	109

Der Levene-Test wird zur Überprüfung der Homogenität der Varianzen eingesetzt und zeigt mit $p = .454$ ein nicht signifikantes Ergebnis. Die Normalverteilung der EMS-ähnlichen Untertests kann ebenfalls angenommen werden. Die Voraussetzungen zur Berechnung des varianzanalytischen Designs können somit als gegeben betrachtet werden.

Tabelle 43: Test der Zwischensubjekteffekte, abhängige Variable: EMS Figuren Lernen

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	<i>F</i>	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	126.473 ^a	23	5.499	1.313	.147	.027
Konstanter Term	18812.631	1	18812.631	4493.728	.000	.808
Geschlecht	3.330	1	3.330	0.795	.373	.001
Stereotype	9.639	3	3.213	0.767	.512	.002
Risikoverhalten	3.656	2	1.828	0.437	.646	.001

Geschl * Stereo	3.302	3	1.101	0.263	.852	.001
Geschl * Risiko	48.092	2	24.046	5.744	.003*	.011
Stereotype * Risiko	16.929	6	2.822	0.674	.671	.004
Geschl * Stereo * Risiko	24.095	6	4.016	0.959	.452	.005
Fehler	4475.282	1069	4.186			
Gesamt	26043.000	1093				
Korr.Gesamtvariation	4601.755	1092				
a. R-Quadrat = ,027 (korrigiertes R-Quadrat = ,007)						

Der Test der Zwischensubjekteffekte zur Überprüfung der $H_1(5)$ für die entsprechenden standardisierten Prüfgröße ergibt mit $F(2,1069) = 5.744$, $p = .003$ eine signifikante Ergebnis. Somit kann eine signifikante Interaktion im EMS- ähnlichen Untertest *Figuren Lernen* in Abhängigkeit von Geschlecht und Risikoverhalten beobachtet werden.

Die Interaktion zeigt an, dass die Differenzen der Leistungen im EMS- ähnlichen Untertest *Figuren Lernen* zwischen den Geschlechtern und in Abhängigkeit von den Risikoinstruktionen unterschiedlich ausfallen. Es ist erkennbar, dass Leistungsdifferenzen in den Risikoinstruktionen in Abhängigkeit vom Geschlecht unterschiedlich sind (siehe Abbildung 24).

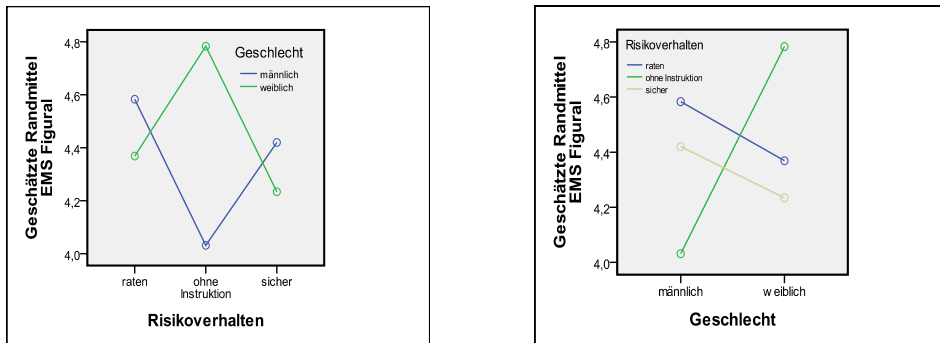


Abbildung 24: Wechselwirkungsdiagramm- Interaktion zwischen Geschlecht und Risikoverhalten

Der Verlauf der Graphen weist auf eine disordinale Interaktion, die aus dem Zusammenwirken der beiden Faktoren Geschlecht und Risikoverhalten zu erklären ist, hin. Sobald der Verlauf der Graphen in beiden Wechselwirkungsdiagrammen nicht gleichsinnig ist, kann von einer disordinalen Interaktion ausgegangen werden (Bortz & Döring, 2006, S 534). Bei einer Interaktion bewirken bestimmte Faktorstufenkombinationen überadditive Effekte, induzieren demnach besonders gute oder schlechte Leistungen, beispielsweise zeigen weibliche Schülerinnen besonders gute Leistungen im Untertest *Figuren Lernen* wenn *keine Instruktion* gegeben wurde. Für männliche Schüler ist die Instruktion zu *Raten* für die Leistung im Untertest *Figuren Lernen* förderlich.

Mädchen zeigen in der Risikoinstruktion *ohne Instruktion* bessere Leistungen als ihre männlichen Mitschüler, während diese Unterschiede in den anderen beiden Instruktionsbedingungen (*sicher* und *raten*) in dieser Form nicht sichtbar werden.

Darüber hinaus ergibt die Prüfung der übrigen Hypothesen keine weiteren signifikanten Haupteffekte und Wechselwirkungen.

Hypothesenblock II.

H₁(8) Es gibt einen signifikanten Unterschied im EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Geschlecht

H₁(9) Es gibt einen signifikanten Unterschied im EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit der Stereotype Threat-Bedingungen (4).

H₁(10) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Risikoverhalten (3).

H₁(11) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Geschlecht und der Stereotype Threat- Bedingungen(4).

H₁(12) Es gibt einen signifikanten Unterschied im EMS- ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Geschlecht und dem Risikoverhalten (3).

H₁(13) Es gibt einen signifikanten Unterschied im EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit der Stereotype Threat-Bedingungen (4) und dem Risikoverhalten (3).

H₁(14) Es gibt eine signifikante Interaktion höherer Ordnung im EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Geschlecht, der Stereotype Threat-Bedingungen (4) und dem Risikoverhalten (3).

In der folgenden Hypothesenprüfung werden wiederum drei unabhängige Variablen bzw. Faktoren, nämlich Geschlecht, Stereotype Threat Bedingungen und das Risikoverhalten bezüglich einer abhängigen Variable, dem EMS-ähnlichen Untertest *Fakten Lernen*, untersucht (siehe Tabelle 44). Aus diesem Grund erfolgt die Hypothesentestung mittels einer univariaten dreifaktoriellen Varianzanalyse.

Tabelle 44: Deskriptiv Statistik * EMS- ähnlicher Untertest Fakten Lernen

		männlich			Weiblich			Gesamt		
Stereo	Risiko	M	SD	n	M	SD	n	M	SD	n
A	Gesamt	4.03	1.846	123	4.47	1.788	170	4.29	1.823	293
B	Gesamt	3.65	1.856	85	4.27	1.770	173	4.06	1.819	257
C	Gesamt	4.12	1.639	112	4.56	1.662	168	4.39	1.664	280
D	Gesamt	3.88	1.678	89	4.39	1.769	174	4.21	1.752	263
Gesamt	Raten	4.11	1.659	145	4.44	1.639	228	4.31	1.653	373
	Ohne	4.04	1.782	131	4.45	1.826	237	4.30	1.818	368
	Sicher	3.67	1.825	133	4.37	1.778	219	4.11	1.825	352
	Gesamt	3.94	1.760	409	4.42	1.748	684	4.24	1.767	1093

Es kann eine Normalverteilung der EMS-ähnlichen Untertest Werte *Fakten Lernen* angenommen werden. Die Prüfung der Homogenität der Varianzen mittels Levene- Test ergibt mit $p = .133$ ein nicht signifikantes Ergebnis.

Tabelle 45: Tests der Zwischensubjekteffekte, abhängige Variable: EMS Fakten

Tests der Zwischensubjekteffekte						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	123.096 ^a	23	5.352	1.742	.017	.036
Konstanter Term	16876.770	1	16876.770	5491.783	.000	.837
Geschlecht	62.150	1	62.150	20.224	.0001**	.019
Stereotype	21.524	3	7.175	2.335	0.072	.007
Risikoverhalten	11.487	2	5.743	1.869	0.155	.003
Geschl * Stereotype	1.637	3	.546	.178	.0912	.000
Geschl * Risiko	7.676	2	3.838	1.249	0.287	.002
Stereo * Risiko	5.089	6	.848	.276	0.948	.002
Geschl * Stereo * Risiko	17.158	6	2.860	.931	0.472	.005
Fehler	3285.139	1069	3.073			
Gesamt	23072.000	1093				
Korrigierte Gesamtvariation	3408.234	1092				
a. R-Quadrat = .036 (korrigiertes R-Quadrat = .015)						

Die Überprüfung der $H_1(8)$ mittels univarianter dreifaktorieller Varianzanalyse (ANOVA) ergibt für das Geschlecht mit der entsprechenden Prüfgröße $F(1,1069) = 20.224$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 45). Es kann ein signifikanter Unterschied im EMS-ähnlichen Untertest *Fakten Lernen* in Abhängigkeit vom Geschlecht angenommen werden. Die entsprechende standardisierte Effektgröße, $d = 0.31$ gibt den Hinweis auf einen mittleren Effekt (Bortz & Döring, 2006, S. 606). Mädchen ($M = 4.42$, $SD = 1.748$) weisen in der Bearbeitung des EMS-ähnlichen Untertests *Fakten Lernen* im Vergleich zu männlichen Schülern höhere Werte auf.

Mit dem Test der Zwischensubjekteffekte zur Überprüfung der $H_1(9)$ für die entsprechende standardisierte Prüfgröße kann mit $F(3, 1069) = 2.335, p = .072$ ein nicht signifikantes Ergebnis beobachtet werden. Es kann kein signifikanter Unterschied im EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit von den Stereotype Threat Bedingungen angenommen werden. Aufgrund des Ergebnisses kann in diesem Fall lediglich von einer tendenziellen Signifikanz ausgegangen werden (siehe Abbildung 25).

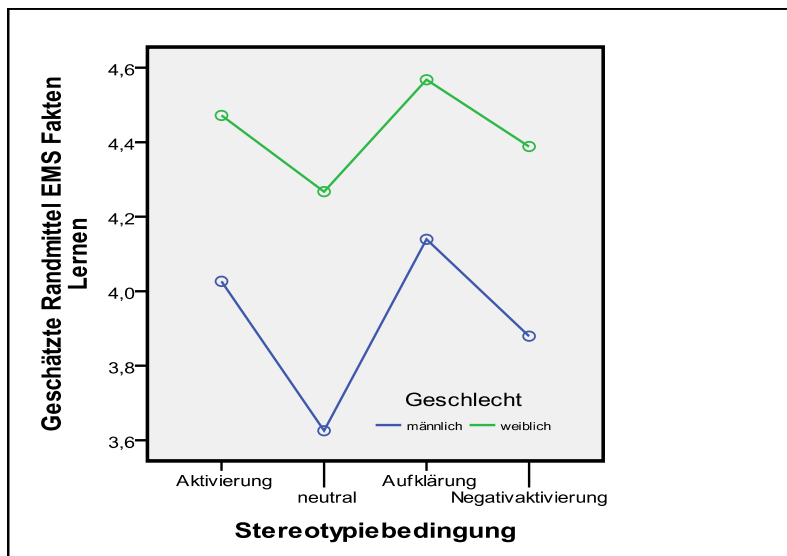


Abbildung 25: Stereotype Threat Bedingungen und Untertest Fakten Lernen

Darüber hinaus können im Rahmen der Hypothesenprüfung keine weiteren signifikanten Haupteffekte und Wechselwirkungen beobachtet werden.

3.4.5.2 Nebenhypothese der EMS- ähnlichen Untertests in

Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den EMS

Im folgenden Abschnitt werden die bedeutenden Nebenhypothesen zu den EMS- ähnlichen Untertests Fakten Lernen und Figuren Lernen beschrieben.

Hypothesenblock III.

H₁ (15) Es gibt signifikante Unterschiede in den Leistungen des EMS-ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den Eingangstest zum Medizinstudium (EMS)

H₁ (16) Es gibt signifikante Unterschiede in den Leistungen des EMS-ähnlichen Untertest Fakten Lernen in Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den Eingangstest zum Medizinstudium (EMS)

Zur Prüfung der relevanten Nebenhypothesen der EMS- ähnlichen Untertests Fakten Lernen und Figuren Lernen wurde eine univariate zweifaktorielle Varianzanalyse (ANOVA) verwendet.

Im Rahmen dieser Berechnungen werden als unabhängige Variable der Wunsch, Medizin zu studieren und die spezifische Vorbereitung auf den Eingangstest zum Medizinstudium (EMS) und als abhängige Variable die beiden EMS- ähnlichen Untertests *Fakten Lernen* bzw. *Figuren Lernen* untersucht.

Insgesamt können 203 (18,3 %) SchülerInnen ermittelt werden, die angaben, sich bereits in irgendeiner Form mit der Vorbereitung auf den EMS beschäftigt zu haben. Hiervon gaben 53 (5,0 %) TestteilnehmerInnen von 1068 gültigen Fällen an, sich bereits speziell mit den EMS-Testmaterial auseinander gesetzt zu haben.

Tabelle 46 zeigt die Häufigkeitsverteilungen im Wunsch, Medizin studieren zu wollen und der spezifischen Vorbereitung auf den EMS-Eingangstest zum Medizinstudium.

Tabelle 46: Vierfelder- Tafel Wunsch Medizin zu studieren und spezifische Vorbereitung auf den EMS

Medizin * EMSTest Kreuztabelle					
			EMSTest		Gesamt
			nein	Ja	
Medizin	nein	Anzahl	860	19	879
		Erwartete Anzahl	835,3	43,7	879,0
		% innerhalb von Medizin	97,8%	2,2%	100,0%
		Standardisierte Residuen	,9	-3,7	
	ja	Anzahl	153	34	187
		Erwartete Anzahl	177,7	9,3	187,0
		% innerhalb von Medizin	81,8%	18,2%	100,0%
		Standardisierte Residuen	-1,9	8,1	
Gesamt		Anzahl	1013	53	1066
		Erwartete Anzahl	1013,0	53,0	1066,0
		% innerhalb von Medizin	95,0%	5,0%	100,0%

Die Berechnung der entsprechenden Prüfgröße ergibt mit $\chi^2(1) = 83.761$, $p \leq .0001$ einen signifikanten Zusammenhang zwischen dem Wunsch, Medizin zu studieren und einer spezifischen Beschäftigung mit dem EMS. Es kann ein Verteilungsunterschied beobachtet werden.

Aufgrund der Gruppengrößen ist eine Berechnung eines varianzanalytischen Designs zulässig.

Ergebnisse *Figuren Lernen* in Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den EMS

Die Überprüfung der Homogenität der Varianzen mittels Levene- Test ergibt mit $p = .621$ ein nicht signifikantes Ergebnis. Des Weiteren kann die Normalverteilung der Daten angenommen werden. Somit sind die

Voraussetzungen für die Berechnung einer univariaten zweifaktoriellen ANOVA als erfüllt zu betrachten.

Die Überprüfung der $H_1(15)$ mittels ANOVA fällt mit der entsprechenden Prüfgröße $F(1,1062) = 9.152$, $p = .003$ für den Faktor Wunsch Medizin zu studieren signifikant aus. Es kann ein signifikanter Unterschied in den Leistungen des EMS-ähnlichen Untertest *Figuren Lernen* in Abhängigkeit vom Wunsch, Medizin zu studieren angenommen werden (siehe Tabelle 47).

Tabelle 47: Test der Zwischensubjekteffekte Untertest Figuren Lernen in Abhängigkeit vom Wunsch Medizin zu studieren und spezifischer Vorbereitung auf den EMS

Tests der Zwischensubjekteffekte					
Abhängige Variable:EMS Figuren					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	<i>F</i>	Sig.
Korrigiertes Modell	70.566 ^a	3	23.522	5.681	.001
Konstanter Term	3855.228	1	3855.228	931.161	.000
Medizin	37.890	1	37.890	9.152	.003
EMSTest	.495	1	.495	.119	.730
Medizin * EMSTest	6.379	1	6.379	1.541	.215
Fehler	4396.935	1062	4.140		
Gesamt	25402.000	1066			
Korrigierte Gesamtvariation	4467.501	1065			
a. R-Quadrat = ,016 (korrigiertes R-Quadrat = ,013)					

Aus Tabelle 48 ist sind Mittelwerte und Standardabweichungen mit Wunsch, Medizin zu studieren (ja/nein) und spezifischer Vorbereitung auf den EMS (ja/nein) in Abhängigkeit des Unterests Figuren Lernen zu entnehmen.

Tabelle 48: Deskriptiv Statistik Wunsch, Medizin zu studieren und spezifische Vorberitung auf den EMS im Untertest Figuren Lernen

Deskriptive Statistiken				
Abhängige Variable:EMS Figuren				
Medizin	EMSTest	<i>M</i>	<i>SW</i>	N
nein	nein	4.33	2.038	860
	ja	4.05	1,.957	19
	Gesamt	4.32	2.036	879
ja	nein	4.87	2.064	153
	ja	5.35	1.840	34
	Gesamt	4.96	2.029	187
Gesamt	nein	4.41	2.050	1013
	ja	4.89	1.968	53
	Gesamt	4.43	2.048	1066

SchülerInnen mit dem Wunsch, Medizin zu studieren ($M = 4.96$, $SD = 2.03$) schneiden im EMS-ähnlichen Untertest Figuren Lernen besser ab als Personen, welche nicht beabsichtigen, ein Medizinstudium ($M = 4.32$, $SD = 2.04$) zu beginnen (siehe Abbildung 26).

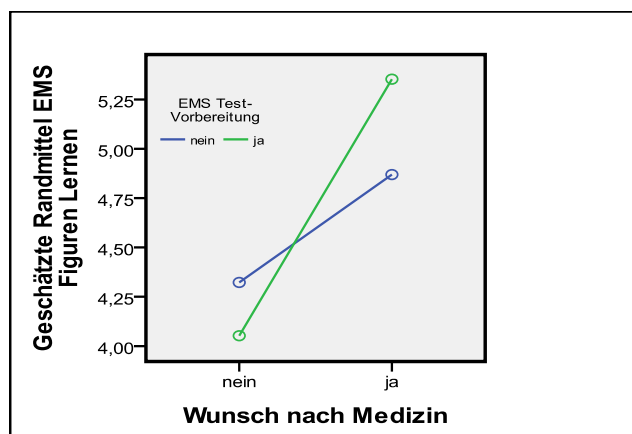


Abbildung 26: Untertest Figuren Lernen in Abhängigkeit Medizinstudium

Des Weiteren können keine Wechselwirkungen im Medizin Wunsch ($p = .215$) in Abhängigkeit von der Vorbereitung mit dem EMS beobachtet werden.

Ergebnisse *Fakten Lernen* in Abhängigkeit vom Wunsch Medizin zu studieren und der spezifischen Vorbereitung auf den EMS

Die Überprüfung der Homogenität der Varianzen für die Nebenhypothese zum EMS-ähnlichen Untertest *Fakten Lernen* mittels Levene-Test, fällt mit $p = .402$ unauffällig aus.

Tabelle 49: Test der Zwischensubjekteffekte Untertest *Fakten Lernen* in Abhängigkeit vom Wunsch Medizin zu studieren und spezifischer Vorbereitung auf den EMS

Tests der Zwischensubjekteffekte					
Abhängige Variable:EMS Fakten					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	33.441 ^a	3	11.147	3.616	.013
Konstanter Term	3747.712	1	3747.712	1215.792	.000
Medizin	.460	1	.460	.149	.699
EMSTest	9.448	1	9.448	3.065	.080
Medizin * EMSTest	3.038	1	3.038	.986	.321
Fehler	3273.644	1062	3.083		
Gesamt	22617.000	1066			
Korrigierte Gesamtvar	3307.085	1065			
a. R-Quadrat = ,010 (korrigiertes R-Quadrat = ,007)					

Im Rahmen des Tests der Zwischensubjekteffekte zur Prüfung der $H_1(16)$ können für den Wunsch, Medizin zu studieren ($p = .699$) für die spezifische Vorbereitung mit dem EMS ($p = .080$) kein signifikanter Effekt und keine Wechselwirkungen ($p = .321$) festgestellt werden.

Tabelle 50: Deskriptiv Statistik Wunsch, Medizin zu studieren und spezifische Vorbereitung auf den EMS im Untertest Fakten Lernen

Deskriptive Statistiken				
Abhängige Variable:EMS Fakten				
Medizin	EMSTest	M	SW	N
nein	nein	4.17	1.777	860
	ja	4.89	1.729	19
	Gesamt	4.19	1.778	879
ja	nein	4.54	1.594	153
	ja	4.74	1.928	34
	Gesamt	4.57	1.656	187
Gesamt	nein	4.23	1.754	1013
	ja	4.79	1.843	53
	Gesamt	4.26	1.762	1066

Es kann angenommen werden, dass keine Unterschiede oder Wechselwirkungen in den Leistungen des EMS-ähnlichen Untertest *Fakten Lernen* in Abhängigkeit vom Wunsch, Medizin zu studieren oder der spezifischen Vorbereitung auf den Eingangstest zum Medizinstudium (EMS) vorhanden sind.

Die vorwiegend nicht signifikanten Ergebnisse deuten darauf hin, dass der Wunsch Medizin zu studieren und die tatsächliche Vorbereitung mit dem Testmaterial im Rahmen des Eingangstest für angehende MedizinerInnen keinen bedeutenden Einfluss auf die beiden abhängigen Variablen EMS-ähnlicher Untertest *Fakten Lernen* und *Figuren Lernen* aufweisen.

3.4.5.3 Hypothesenbezogenen Ergebnisse des Game of Dice Task (GDT)

Im Folgenden Abschnitt sollen die relevanten Nebenhypothesen zum GDT überprüft und beschrieben werden.

Hypothesenblock IV.

H₁(17) Es gibt einen signifikanten Unterschied im Risikoverhalten, operationalisiert durch den Game of Dice Task (GDT), in Abhängigkeit vom Geschlecht

H₁(18) Es gibt einen signifikanten Unterschied im Risikoverhalten, operationalisiert durch den Game of Dice Task (GDT), in Abhängigkeit der Risikoinstruktion

H₁(19) Es ist eine signifikante Wechselwirkung im Risikoverhalten, operationalisiert durch den Game of Dice Task (GDT) in Abhängigkeit vom Geschlecht und der Risikoinstruktion zu beobachten

Die Überprüfung der interessierenden Hypothesen in Bezug auf den GDT erfolgte mit einer univariaten zweifaktoriellen Varianzanalyse (ANOVA). Es wurden als unabhängige Variable das Geschlecht und die Risikoinstruktion und als abhängige Variable das Risikoverhalten, operationalisiert durch den GDT berücksichtigt.

Aus zeitökonomischen Gründen konnte der Game of Dice Task (GDT) nicht allen Personen vorgegeben werden, insgesamt können in die Analysen 471 TeilnehmerInnen miteinbezogen werden.

Die Prüfung der Homogenität der Varianzen der Daten erfolgte mittels Levene- Test. Dieser ergab mit $p = .922$ ein nicht signifikantes Ergebnis. Die Normalverteilung der Daten je Gruppe kann angenommen werden. Im Bearbeiten der Würfelspielaufgaben im Rahmen des GDT weisen niedrige GDT-Nettoscores (NS) auf riskantes Verhalten hin, höhere Werte hingegen zeigen, dass die TeilnehmerInnen eher risikoaversiv handeln und entscheiden. Tabelle xy verdeutlicht dies anhand der Mittelwerte, aufgeteilt in die drei Risikobedingungen (Raten, Ohne Instruktion, Sicher)und

nach dem Geschlecht. Je niedriger der Wert, desto riskanter agiert die Person im Bearbeiten der Würfelspielaufgaben im GDT.

Tabelle 51: Kennwerte Bedingungen Risikoinstruktion

Deskriptive Statistiken				
Abhängige Variable: NS-GDT				
Geschlecht	Risikoverhalten	<i>M</i>	<i>SD</i>	N
männlich	raten	1.25	9.712	61
	ohne Instruktion	4.64	9.737	61
	sicher	3.85	1.,036	61
	Gesamt	3.25	9.883	183
weiblich	raten	3.66	9.579	96
	ohne Instruktion	5.66	9.079	102
	sicher	5.63	9.976	90
	Gesamt	4.98	9.546	288
Gesamt	raten	2.72	9.672	157
	ohne Instruktion	5.28	9.314	163
	sicher	4.91	1.,005	151
	Gesamt	4.31	9.705	471

Die Tabelle zeigt zunächst, dass die Mittelwerte in der Risikoinstruktionsbedingung *raten* ($M = 2.72$, $SD = 9.672$) am niedrigsten sind. Das könnte darauf hindeuten, dass jene Personen, welche zu Beginn zum Raten aufgefordert wurden, auch im Rahmen der Bearbeitung des GDT das risikoreichste Entscheidungsverhalten zeigten. Die Ergebnisse der Hypothesenprüfung werden in Tabelle 52 dargestellt:

Tabelle 52: Test der Zwischensubjekteffekte, abhängige Variable: NS-GDT

Quelle	AQS vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	975.765 ^a	5	195.153	2.096	.065	.022
Konstanter Term	756.,904	1	7567.904	8.,290	.0001	.149
Geschlecht	336.981	1	336.981	3.620	.058	.008
Risikoverhalten	633.373	2	316.687	3.402	.034*	.014
Geschl * Risiko	36.736	2	18.368	0.197	.821	.001
Fehler	43290.596	465	93.098			
Gesamt	53007.000	471				
Korrigierte Gesamtvar	44266.361	470				
a. R-Quadrat = .022 (korrigiertes R-Quadrat = .012)						

Der Test der Zwischensubjekteffekte zur Überprüfung der $H_1(16)$ ergibt mit der entsprechenden standardisierten Prüfgröße $F(2, 465) = 3.402$, $p = .034$ ein signifikantes Ergebnis. Somit kann ein signifikanter Unterschied im Risikoverhalten, operationalisiert mit dem GDT, in Abhängigkeit von der Risikoinstruktion angenommen werden.

Die im Rahmen der Varianzanalyse durchgeführte Prüfung der $H_1(15)$ ergibt mit $F(1, 465) = 3.620$, $p = .058$ ein nicht signifikantes Ergebnis. Es kann lediglich ein tendenzieller Niveauunterschied in Abhängigkeit vom Geschlecht angenommen werden. Die Mittelwerte weisen darauf hin, dass männliche Schüler ($M = 3.25$, $SD = 9.883$) im Gegensatz zu Mädchen ($M = 4.98$, $SD = 9.546$) ein etwas riskanteres Entscheidungsverhalten zeigen.

Ebenso ergibt die Prüfung der $H_1(17)$ mittels Berechnung der ANOVA mit der entsprechenden Prüfgröße $F(2, 465) = 0.197$, $p = .821$ ein nicht signifikantes Ergebnis. Es kann keine signifikante Wechselwirkung im Risikoverhalten, operationalisiert mittels GDT, in Abhängigkeit vom Geschlecht und der Risikoinstruktion angenommen werden.

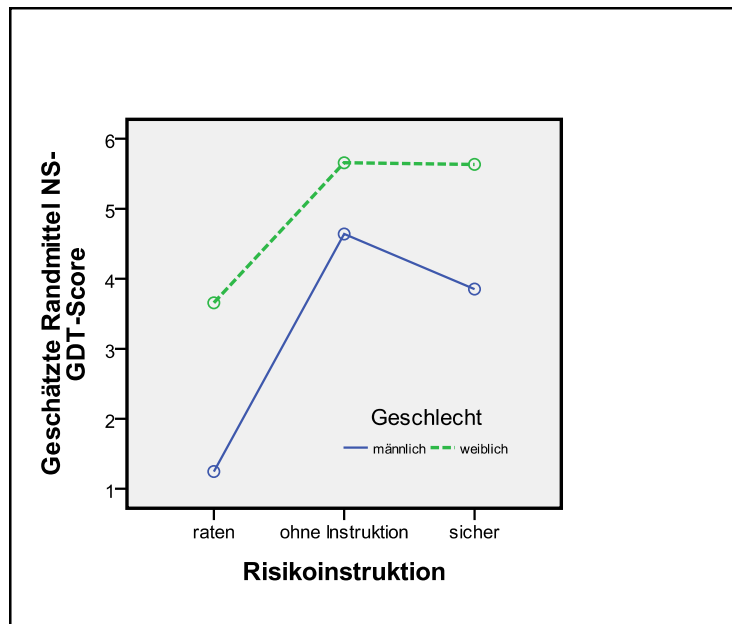


Abbildung 27: NS-GDT Score in Abhängigkeit vom Geschlecht und Risikoinstruktion

Die Abbildung verdeutlicht, dass männliche Schüler im Rahmen der Würfelspielaufgaben des GDT tendenziell risikoreicher entscheiden als Mädchen. In der Rateinstruktion scheinen diese Unterschiede am stärksten ausgeprägt zu sein.

Aufgrund der Vorannahme, dass eine Rateinstruktion riskantes Entscheidungsverhalten fördert, werden lineare Kontraste berechnet, um die möglichen Unterschiede im Risikoverhalten in Abhängigkeit der jeweiligen Risikoinstruktion zu prüfen.

Die Ergebnisse zeigen, dass sich die Risikobedingung *raten* signifikant von den anderen beiden Risikoinstruktionen unterscheidet (Vergleich *raten* gegen *sicher*, $p = .042$ und Vergleich *raten* gegen *ohne Instruktion*, $p = .016$).

Aufgrund der Ergebnisse kann man davon ausgehen, dass die TeilnehmerInnen, sofern sie unter der Rateinstruktion arbeiten mussten,

auch im GDT riskanteres Entscheidungsverhalten, im Gegensatz zu den anderen beiden Risikoinstruktionen (ohne Instruktion, sicher), zeigten.

4 DISKUSSION UND AUSBLICK

Im folgenden Teil werden die Ergebnisse der vorliegenden empirischen Arbeit im Überblick zusammengefasst dargestellt. Des Weiteren soll eine Verbindung zwischen den im Theorieteil beschriebenen Studien und der aktuell durchgeführten Untersuchung hergestellt werden. Abschließend werden Stärken und Schwächen dieser Untersuchung in Bezug auf deren Durchführung dargestellt, um Anregungen und Perspektiven für mögliche weiterführende Untersuchungen aufzuzeigen.

Das Hauptziel der vorliegenden empirischen Arbeit war die Untersuchung von Geschlechtsunterschieden in Bezug auf die Bearbeitung EMS-ähnlicher Aufgaben. Hauptaugenmerk wurde auf jene Bereiche gelegt, welche Gedächtnisleistungen erfassen und in Abhängigkeit sozialer Einflussfaktoren wie Stereotype Threat und Risikoverhalten.

In der Tat ist die Untersuchung von Geschlechtsunterschieden, vorallem in Bezug auf kognitive Leistungsbereiche eine häufig diskutierte und nach wie vor hochaktuelle Thematik. Dies wurde ebenfalls in dieser Arbeit aufgezeigt, die sich mit dem Bereich der Studieneignung zum Medizinstudium, welche mittels dem EMS erfasst wird, beschäftigt hat.

Wie aus einer Untersuchung von Spiel et al. (2008) hervorgeht unterscheiden sich die Geschlechter wesentlich in Bezug auf ihre Leistungen im Eignungstest für das Medizinstudium. Frauen schneiden in sieben (Schlauchfiguren, Medizinisch-naturwissenschaftliches Grundverständnis, Planen und Organsieren, Diagramme und Tabellen, Konzentriertes und sorgfältiges Arbeiten, Textverständnis und Quantitative und Formale Probleme) von den insgesamt zehn Untertests bedeutend schlechter ab als ihre männlichen Mitstreiter.

Die Tatsache, dass letztlich mehr Frauen als Männer ein Medizinstudium erfolgreich abschliessen und sich somit die Leistungsunterschiede während dem Medizinstudium zur Gänze auflösen (Spiel et. al., 2008), lässt die Vermutung zu, dass das schlechtere Abschneiden im Eignungstest für das Medizinstudium (EMS) auf andere Gegebenheiten und Ursachen, wie beispielsweise die Testvorgabe und soziale Faktoren, wie Geschlechterrollen bzw. Geschlechterstereotype deren Aktivierung durch jene Testvorgabe begünstigt wird, begründet ist und vorhandene Geschlechtsunterschiede zugunsten der Männer nicht vollständig auf die Unterschiede im kognitiven Leistungsvermögen zurückzuführen sind.

Der Wissenstest, der in Anlehnung an die Untersuchung von Halpern und Tan (2001) vorgegeben wurde, diente in der vorliegenden Untersuchung zur Überprüfung der Wirksamkeit der experimentellen Variationen von den Stereotype Threat Bedingungen (*Aktivierung, Neutral, Aufklärung und Negativaktivierung*). Im Rahmen dazu wurde jeweils eine Fremdeinschätzungsaufgabe und eine Selbsteinschätzungsaufgabe in Bezug auf den Bereich Lernen und Gedächtnis vorgegeben. Die Ergebnisse machen deutlich, dass die Versuchspersonen durchwegs den jeweiligen Stereotype Threat Bedingungen entsprechend geantwortet haben. Beispielsweise wurden den Frauen in der Aktivierungsbedingung gemäß dem Stereotyp, sowohl von den weiblichen als auch von den männlichen Versuchspersonen bessere Gedächtnisleistungen zugesprochen, im Gegensatz dazu wurden in der Stereotype Threat Bedingung, *Negativaktivierung* den Männern bessere Gedächtnisleistungen zugesprochen. In einer Vielzahl von Untersuchungen konnte ermittelt werden, dass sich die Aktivierung von Stereotype Threat vor der Bearbeitung von kognitiven Aufgaben, ungünstig auf die anschliessenden Leistungen von Frauen auswirkt (Keller & Dauenheimer, 2003; Marx & Stapel, 2006; Spencer et.al. 1999;).

Ein weiteres Indiz für den Einfluss von Stereotype Threat auf die kognitive Leistungsfähigkeit geben auch die Ergebnisse des Wissenstest in Bezug auf die Stereotype Threat Bedingung, *Aufklärung*. In jener Bedingung

wurden die TestteilnehmerInnen über das Phänomen Stereotype Threat aufgeklärt und jene Information hatten maßgeblichen Einfluss auf die Einschätzungen der TestteilnehmerInnen. Demzufolge wurden, anders als in der Aktivierungs- bzw. Negativaktivierungsbedingung, in der Aufklärungsbedingung sowohl den Männern als auch den Frauen gleich gute Gedächtnisleistungen zugesprochen. Zu ähnlichen Ergebnissen kamen in ihrer Untersuchung auch Johns et. al. (2005), in der gezeigt werden konnte, dass sich durch die Aufklärung des Phänomens Stereotype Threat Leistungsunterschiede nahezu auflösen und Frauen, wenn man sie über das besagte Phänomen aufklärt, bessere Leistungen aufweisen, als Frauen bei denen Stereotype Threat aktiviert wurde und sich ihre Leistungen nicht von jenen der untersuchten Männer unterschieden.

Wie bereits erwähnt war das Ziel der vorliegenden Untersuchung zu ermitteln ob sich Männer und Frauen beim Bearbeiten EMS- ähnlicher Aufgaben, welche die Merkfähigkeit erfassen (Fakten- und Figuren Lernen) signifikant voneinander unterscheiden und dies unter Berücksichtigung möglicher wechselseitiger Einflüsse durch Stereotype Threat und Risikoverhalten.

Für den EMS- ähnlichen Untertest *Fakten Lernen* konnte in Bezug auf die Leistungen ein signifikanter Unterschied in Abhängigkeit vom Geschlecht festgestellt werden (H_1 8). Im Vergleich zeigte sich, dass Mädchen ($M = 4.42$; $SD = 1.748$) im Fakten Lernen signifikant bessere Leistungen zeigten als Männer ($M = 3.94$; $SD = 1.760$). Einige in der Literatur beschriebenen Untersuchungen kamen zu analogen Ergebnissen, so zeigte beispielsweise eine Untersuchung von Kramer et.al. (1997) in Bezug auf Gedächtnisleistung und verbalem Lernen, dass die Wiedergabeleistung in einem verbalen Lerntest, sowohl in der freien Wiedergabe Bedingung, als auch in der Wiedererkennungslleistung zwanzig Minuten nach Vorgabe zugunsten der Frauen beobachtet werden konnte. Zu ähnlichen Ergebnissen, bei der eine Überlegenheit von Frauen hinsichtlich Reproduktions- als auch Wiedererkennungslleistung dreißig

Minuten nach Vorgabe eines Lerntests resultierte, kam eine Untersuchung von West et.al. (1992). Diese sahen die Überlegenheit der Frauen unter anderem jedoch im Geschlechterstereotypen Testmaterial (Lebensmittellisten und Vor- Nachnamen- Lernen) begründet und vermuteten Frauen hätten möglicherweise mehr Erfahrung mit Lebensmittellisten und wären eher bemüht gesellschaftlichen Stereotypen zu entsprechen und beispielsweise als sozial kompetenter zu gelten als Männer und seien aus diesem Grund motivierter bei der Bearbeitung im Untertest Namen-Lernen.

Die Ergebnisse der vorliegenden Untersuchung für den Untertest Figuren Lernen zeigen, dass keine signifikanten Unterschiede im Geschlecht beobachtet werden konnten (H_{11}). Analoge Ergebnisse konnten auch in Untersuchungen von Ionescu (2000, 2002) ermittelt werden, der keine Geschlechterunterschiede in Bezug auf visuelle Merkfähigkeitsleistungen feststellte. In einer Studie von Choi und L'Hirondelle (2005) im Rahmen geschlechtsspezifischer Unterschiede in Bezug auf die Lokalisation von Objekten konnte beobachtet werden, dass der Kontext einen wesentlichen Einfluss auf die Richtung der Geschlechtsunterschiede hat und eine Überlegenheit der Frauen nur in jener Bedingung festzustellen war, in der konkrete Objekte dargeboten wurden, nicht aber bei der Vorgabe abstrakter Objekte. Der Vorteil der Frauen hinsichtlich der Lokalisation von konkreten Objekten sehen die Autoren unter anderem in der Überlegenheit der Frauen in verbalen Gedächtnisleistungen begründet (Choi & L'Hirondelle, 2005).

Für beide EMS- ähnlichen Untertests (Fakten- und Figuren Lernen) zur Erfassung der Merkfähigkeit konnten in Abhängigkeit der Stereotype Threat Bedingungen keine bzw. nur tendenziell signifikante Unterschiede beobachtet werden (H_{12} ; H_{19}). Betrachtet man weiters die Ergebnisse der beiden EMS- ähnlichen Aufgaben in Abhängigkeit der Stereotype Threat Bedingungen und dem Geschlecht, so konnten auch diesbezüglich keine signifikanten Ergebnisse ermittelt werden (H_{14} ; H_{111}). Im Gegensatz zu den Ergebnissen der vorliegenden Untersuchung konnten

beispielsweise Schmader und Johns (2003) feststellen, dass sich bei Vorgabe eines Gedächtnistest unter einer Stereotype Threat aktivierenden Bedingung die Gedächtnisleistung bei Frauen vermindert zeigte. Im Vergleich zu den weiblichen Versuchsteilnehmern ohne Stereotype Threat Instruktion und den männlichen Versuchsteilnehmern zeigten jene Frauen schlechtere Gedächtnisleistungen. Colley et al. (2002) fanden ebenfalls heraus, dass Frauen verminderte Gedächtnisleistungen zeigen, wenn man ihnen zuvor mitteilte, dass Männer in jenem Bereich oftmals besonders gute Leistungen zeigten. Bei Männern hingegen hatten die unterschiedlichen Instruktionsbedingungen keinen Einfluss auf die Gedächtnisleistung im darauffolgenden Test.

Ein weiteres interessantes Ergebnis der aktuell durchgeführten Untersuchung zeigte sich für den EMS- ähnlichen Untertest Figuren Lernen, für den eine signifikante Wechselwirkung in Abhängigkeit vom Geschlecht und der Risikoinstruktion beobachtet werden konnte ($H_1 5$). Für Frauen scheint die Risikoinstruktion *Ohne Instruktion* förderlicher für die Leistung im Figuren Lernen zu sein, als für Männer, welche in dieser Risikoinstruktionsbedingung schlechtere Leistungen aufweisen. Im Gegensatz dazu scheint sich die Risikoinstruktion Raten besonders förderlich auf die Leistung der Männer auszuwirken.

Folgendes Bild spiegelt auch die Ergebnisse des Game of Dice Task (GDT), der in der vorliegenden Untersuchung zusätzlich zu den unterschiedlichen Risikoinstruktionen zur Erfassung des Risikoverhaltens im Anschluss an die EMS- ähnlichen Untertests vorgegeben wurde, wider. Die Ergebnisse zeigen einen signifikanten Unterschied im Risikoverhalten, operationalisiert mittels GDT in Abhängigkeit der Risikoinstruktion (H_{18}) und ein knapp signifikantes Ergebnis für das Geschlecht (H_{17}). Männer ($M = 3.25$; $SD = 9.883$) weisen auch im Bearbeiten des GDT im Vergleich zu Frauen ($M = 4.98$; $SD = 9.546$) ein etwas riskanteres Entscheidungsverhalten auf. Zudem scheint vorallem die Risikoinstruktion Raten ein risikoreiches Entscheiden zu begünstigen.

Werden nun die Ergebnisse der vorliegenden Arbeit mit jenen in der Literatur verglichen, so zeigt sich in den unterschiedlichsten Bereichen, dass sich die Geschlechter in Bezug auf Riskoverhalten unterscheiden und Männer eher zu riskanterem Verhalten neigen als Frauen (Byrnes et al., 1999; Eckel & Grossmann, 2008; Levin et al., 1988). Zu ähnlichen Ergebnissen gelangen auch Ben-Shakar & Sinai (1991), welche sich in ihrer Untersuchung mit dem Zusammenhang zwischen der Tendenz zu raten und dem Geschlecht bei der Bearbeitung von Multiple Choice Aufgaben beschäftigten. Ben-Shakar und Sinai konnten zeigen, dass Frauen mehr als Männer dazu neigen Items auszulassen, wenn sie sich hinsichtlich der Antwort nicht sicher waren und diese Geschlechtsunterschiede auch nicht verschwanden, wenn die VersuchsteilnehmerInnen vor Bearbeitung zum Raten aufgefordert wurden.

Zusätzlich von Interesse im Rahmen der aktuell durchgeführten Untersuchung war, ob eine vorangegegangene spezifische Beschäftigung mit dem Testmaterial zum Eignungstest für das Medizinstudium (EMS) einen Einfluss auf die Testergebnisse in den Untertests Fakten- und Figuren lernen zeigen bzw. diese sogar verbessern (H_{15} ; H_{16}). In Bezug auf diese Fragestellungen konnten keine signifikanten Ergebnisse beobachtet werden. Somit kann angenommen werden, dass die Beschäftigung mit dem Testmaterial zum EMS im Rahmen dieser Untersuchung als Störvariable mit einiger Sicherheit ausgeschlossen werden kann. Unklar bleibt jedoch wie intensiv das Ausmaß der Beschäftigung mit dem EMS-Testmaterial ausfiel, da es in der vorliegenden Untersuchung nicht erhoben werden konnte. In Bezug auf die Trainierbarkeit des EMS wurde laut Mallinger et al., (2009) in der Entwicklung der Aufgaben darauf geachtet, dass bei der Bearbeitung der Aufgaben kein spezifisches Vorwissen benötigt wird, um die Trainierbarkeit der Aufgaben und deren Lösungen gering zu halten (Mallinger et al., 2009).

Die Reliabilitätsanalysen der EMS-ähnlichen Untertests die für die aktuell durchgeführte Untersuchung mittels Cronbach's Alpha ermittelt wurden,

zeigen bis auf die beiden Untertests, *Figuren Lernen* und *Schlauchfiguren* relativ geringe und unbefriedigende Reliabilitäten. Das Vorhandensein eines niedrigen Cronbach's Alpha einer Skala lässt vermuten, dass im Rahmen der Vorgabe dieser Untertests nicht nur eine einzige Fähigkeit erfasst wird und somit die Skalierung als unzufriedenstellend angenommen werden kann. Das Nebengütekriterium Skalierung ist dann erfüllt, wenn die resultierenden Werte eines Tests die empirischen Verhaltensrelationen adäquat abbilden (Kubinger, 2006, S. 79). Die vorhandenen Geschlechtsunterschiede, welche in der vorliegenden Untersuchung in einigen EMS-ähnlichen Untertests zugunsten der männlichen Testteilnehmer ermittelt werden konnten, sollte aufgrund der nicht hinreichend vorhandenen Skalierung der Skalen relativ gesehen werden. Die Vermutung liegt nahe, dass nicht nur die intendierte Fähigkeit erfasst wird, Männer schneiden somit eventuell deswegen besser ab, weil sie mit der Speed-Komponente im Test besser umgehen können als Frauen.

Zusammenfassend kann gesagt werden, dass die vorliegenden Untersuchung, abgesehen von einigen Aspekten, einen guten Ansatz für weiterführende Untersuchungen darstellt.

Es gibt jedoch, abgesehen von der erhobenen Selbstwirksamkeitserwartung und Leistungsmotivation, deren Ergebnisse im Detail in den Arbeiten von Gabriele Hangl und Anita Teufl nachzulesen sind, einige potenzielle Persönlichkeitseigenschaften, wie beispielsweise Ängstlichkeit und physiologische Korrelate, wie der Hormonstatus, welche im Rahmen dieser Untersuchung aufgrund beschränkter Ressourcen nicht in die Analysen miteinbezogen werden konnten. Dennoch wäre es für weiterführende Untersuchungen zu diesem Themengebiet von großem Interesse inwieweit diese Konstrukte einen Einfluss auf die Leistungen im Eignungstest zum Medizinstudium haben.

5 ZUSAMMENFASSUNG DER ERGEBNISSE

Im Anschluss werden die Ergebnisse der vorliegenden Untersuchung zusammenfassend unter Einbeziehung und Berücksichtigung der formulierten Hypothesen und durchgeführten statistischen Analysen dargestellt.

Im Rahmen dieser größer angelegten Forschungsarbeit nahmen insgesamt 1093 SchülerInnen der 12. Schulstufe einer Allgemein Bildenden Höheren Schule (AHS), die zum Zeitpunkt der Testung im Alter zwischen 16 und 21 Jahren waren, teil.

Als Hauptziel dieser Untersuchung galt es verschiedene Ansätze und Erklärungen für das unterschiedliche Abschneiden von Männern und Frauen im Eignungstest für das Medizinstudium (EMS) zu finden und dies unter Berücksichtigung von sozialen Einflussfaktoren, wie Stereotype Threat, Risikoverhalten und Theory of Mind, als auch Mediatorvariablen wie Leistungsmotivation und Selbstwirksamkeitserwartung.

Während man die Ergebnisse hinsichtlich des Einflusses der Selbstwirksamkeitserwartung und Leistungsmotivation im Detail in den Arbeiten von Gabriele Hangl und Anita Teufl findet, sind Forschungsergebnisse in Bezug auf die Theory of Mind der Arbeit von Katrin Anzirk zu entnehmen und detaillierte Ergebnisse in Bezug auf Geschlechtsunterschiede im Rahmen von Schlussfolgerndem Denken in der Arbeit von Silvia Hameseder nachzulesen.

Die vorliegende Forschungsarbeit konzentrierte sich im speziellen auf etwaige Geschlechtsspezifische Unterschiede in EMS-ähnlichen Untertests, welche die Merkfähigkeit erfassen unter Berücksichtigung von Stereotype Threat und dem Risikoverhalten.

Die zuvor formulierten und interessierenden Haupt- und Nebenhypothesen wurden mittels univariaten zwei- bzw. dreifaktorieller Varianzanalysen (ANOVA) berechnet.

Auf der Basis der Studie von Halpern und Tan (2001) wurde im Rahmen dieser Untersuchung ein Wissenstest zur Überprüfung der experimentellen Variationen der Stereotype Threat Bedingungen, vorgegeben. Durch die Bearbeitung einer Fremd- bzw. Selbsteinschätzungsaufgabe zum Themenbereich Lernen und Gedächtnis konnte beobachtet werden, dass die TestteilnehmerInnen übereinstimmend mit der jeweiligen Stereotype Threat Bedingung antworteten. Dementsprechend wurden analog zur Stereotype Threat *Aktivierungsbedingung* Frauen bessere Gedächtnisleistungen zugewiesen, wohingegen die Stereotype Threat Bedingung *Aufklärung* dazu führte, dass beiden Geschlechtern gleich gute Gedächtnisleistungen zugesprochen wurden.

Die Ergebnisse in Bezug auf den EMS-ähnlichen Untertest Fakten Lernen ergaben in Abhängigkeit vom Geschlecht (H_{18}) einen signifikanten Unterschied. Frauen weisen im Untertest Fakten Lernen bessere Leistungen auf als Männer. Die entsprechende standardisierte Effektgröße mit $d = 0.31$ weist auf einen mittleren Effekt hin (Bortz & Döring, 2006, S. 606). Im Gegensatz dazu zeigten sich hypothesenkonträr für den EMS-ähnlichen Untertest Figuren Lernen in Abhängigkeit vom Geschlecht (H_{11}) keine signifikanten Unterschiede.

Die Ergebnisse für die EMS-ähnlichen Untertests Fakten- und Figuren Lernen zeigen in Abhängigkeit der unterschiedlichen Stereotype Threat Bedingungen keine bzw. nur tendenziell signifikante Unterschiede (H_{12} ; H_{19}). Zudem konnten ebenso hypothesenkonträr für beide EMS-ähnlichen Untertests keine signifikanten Unterschiede in Abhängigkeit der Stereotype Threat Bedingungen und dem Geschlecht (H_{14} ; H_{111}) beobachtet werden.

In Bezug auf die Leistungen im EMS-ähnlichen Untertest Figuren Lernen konnte eine signifikante Wechselwirkung zwischen den Geschlechtern und in Abhängigkeit der unterschiedlichen Risikoinstruktionen festgestellt werden (H_{15}). Beobachtet werden konnte, dass die

geschlechtsspezifischen Leistungsdifferenzen im Untertest Figuren Lernen abhängig von der jeweiligen Risikoinstruktion unter der gearbeitet wurde, unterschiedlich ausfielen. Bei den männlichen Versuchsteilnehmern scheint sich die Risikoinstruktion *Raten* besonders günstig auf die Leistung im Figuren Lernen auszuwirken, hingegen bei den weiblichen Versuchsteilnehmern wirkt die Risikoinstruktion *Ohne Instruktion* einer besseren Leistung entgegen.

Des Weiteren war von Interesse ob eine spezifische Beschäftigung mit dem EMS-Testmaterial einen Einfluss auf die Leistungen in den EMS-ähnlichen Untertests Fakten- und Figuren Lernen hat (H_{15} ; H_{16}). Hypothesenkonträr konnten keine signifikanten Unterschiede, mittels zweifaktorieller Varianzanalyse in den Leistungen im Fakten- und Figuren Lernen in Abhängigkeit einer vorangegangenen Beschäftigung mit dem Testmaterial zum EMS beobachtet werden. Somit scheint eine spezifische Vorbereitung mit dem EMS-Testmaterial keinen bedeutenden Einfluss auf die Leistungen im Untertest Figuren und Fakten Lernen zu haben.

Die Ergebnisse im Rahmen der Vorgabe des Game of Dice Task (Brand et al., 2005) zur Erfassung des Risikoverhaltens im Anschluss an die EMS-ähnlichen Aufgaben mittels univariater zweifaktorieller Varianzanalyse ergibt hypothesenkonform einen signifikanten Unterschied im Risikoverhalten, operationalisiert mittels GDT in Abhängigkeit der Risikoinstruktionen (H_{18}) und ein tendenziell signifikantes Ergebnis in Abhängigkeit vom Geschlecht (H_{17}). Demnach scheinen auch Männer im Risikoverhalten, erfasst mittels GDT tendenziell zu riskanterem Entscheidungsverhalten zu neigen als Frauen. Die Vorannahme, dass durch die Rateinstruktion zu Beginn der EMS-ähnlichen Aufgaben riskantes Entscheidungsverhalten begünstigt wird konnte hypothesenkonform mittels Linearer Kontraste ermittelt werden, somit unterscheidet sich die Risikoinstruktion *raten* signifikant von den anderen beiden Risikobedingungen. Jene VersuchsteilnehmerInnen, sofern sie unter der Risikobedingung *Raten* gearbeitet hatten, zeigten auch im GDT riskanteres Entscheidungsverhalten.

6 ABSTRACT

Fragestellung

Ziel der vorliegenden empirischen Studie war es geschlechtsspezifische Unterschiede im Rahmen EMS-ähnlicher Aufgaben, welche im speziellen Gedächtnisleistungen betreffen, zu ermitteln. Zudem galt es herauszufinden ob und inwieweit soziale Einflussfaktoren Stereotype Threat und Risikoverhalten in Bezug auf Unterschiede zwischen Männern und Frauen im Bearbeiten EMS-ähnlicher Aufgaben, beteiligt sind.

Methode

Die Stichprobe umfasste insgesamt 1093 SchülerInnen der 12. Schulstufe einer Allgemein Bildenden Höheren Schule in Wien, Niederösterreich und dem Burgenland. Den männlichen und weiblichen VersuchsteilnehmerInnen, welche zum Zeitpunkt der Testung im Alter zwischen 16 und 21 Jahren alt waren, wurde vor der Bearbeitung der EMS-ähnlichen Aufgaben, ein Wissenstest vorgegeben, der in Anlehnung an die Studie von Halpern und Tan (2001) in der vorliegenden Untersuchung zur Überprüfung der experimentellen Variationen der Stereotype Threat Bedingungen, eingesetzt wurde. Im Anschluss daran wurden von den SchülerInnen unter verschiedenen Stereotype Threat, als auch Risikoinstruktionsbedingungen EMS-ähnliche Aufgaben bearbeitet. Zudem wurden zusätzlich zu den unterschiedlichen Risikoinstruktionen Würfelspielaufgaben mittels Computer im Rahmen der Vorgabe des Game of Dice Task (Brand et al., 2005) zur Erfassung des Risikoverhaltens vorgegeben.

Ergebnisse

Die Ergebnisse für den EMS-ähnlichen Untertest *Fakten Lernen* zeigen einen signifikanten Unterschied in Abhängigkeit vom Geschlecht, zugunsten der weiblichen TestteilnehmerInnen. Frauen weisen im Untertest *Fakten Lernen* bessere Leistungen auf als Männer. Im Gegensatz dazu konnten für den Untertest *Figuren Lernen* keine signifikanten Unterschiede zwischen Männern und Frauen beobachtet

werden. Für beide EMS-ähnlichen Untertests, welche Gedächtnisleistungen erfassen konnten in Abhängigkeit der verschiedenen Stereotype Threat Bedingungen und dem Geschlecht keine bzw. nur tendenziell signifikante Unterschiede ermittelt werden. Für den EMS-ähnlichen Untertest Figuren Lernen ergab sich eine signifikante Interaktion für das Geschlecht und der Risikoinstruktion. Des Weiteren konnte ermittelt werden, dass eine spezifische vorangegangene Vorbereitung mit dem EMS- Testmaterial keinen signifikanten Einfluss auf die Leistungen in den Untertests Fakten- und Figuren Lernen hatte. Die Ergebnisse des Game of Dice Task (GDT) zeigen signifikante Unterschiede im Risikoverhalten in Abhängigkeit der jeweiligen Risikoinstruktion.

7 LITERATURVERZEICHNIS

- Ambady, N., Paik, S.K., Steele, J., Owen-Smith, A. & Mitchell, J.P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, 40, 401-408.
- Aronson, E., Wilson, T.D. & Akert, R.M. (2008). *Sozialpsychologie* (6. aktualisierte Aufl.). München: Pearson Studium.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2008). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (12. überarbeitete Aufl.). Berlin: Springer.
- Balch, W. R. & Lewis, B. S. (1996). Music-dependent memory: The roles of tempo change and mood meditation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (6), 1354-1363.
- 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.
- Bierhoff, H. W. (2000). *Sozialpsychologie. Ein Lehrbuch*. Stuttgart: Kohlhammer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. überarbeitete Aufl.). Heidelberg: Springer.
- Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J. & Markowitsch, H. 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., Kalbe, E., Labudda, K., Fujiwara, E., Kessler, J. & Markowitsch, H.J. (2005). Decision-making impairments in patients with pathological gambling. *Psychiatry Research*, 133, 91-99.

- Brand, M., Labudda, K., Kalbe, E., Hilker, R., Emmans, D., Fuchs, G., Kessler, J. & Markowitsch, H.J. (2004). Decision-making impairments in patients with Parkinson's disease. *Behavioural Neurology*, 15, 77-85.
- Brining, M. F. (1995). Does mediation systematically disadvantage women? *William and Mary Journal of Women and the Law*, 2, 1-34.
- Bühl, A. & Zöfel, P. (2000). *SPSS Version 9: Einführung in die moderne Datenanalyse unter Windows*. München: Addison-Wesely.
- Byrnes, J. P., Miller, D.C. & Schafer, W.D. (1999). Gender differences in risk taking: A Meta-Analysis. *Psychological Bulletin*, 125 (3), 367-383.
- Carr, P. B. & Steele, C.M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, 45, 853-859.
- Colley, A., Ball, J., Kirby, N., Harvey, R. & Vingelen, I. (2002). Gender-linked differences in everyday memory performance: effort makes the difference. *Sex Roles* 47, 577-582.
- Eckel, C. C. & Grossman, P. J. (2008). Men, women and risk aversion: experimental evidence. In C. Plott & V. Smith (Eds.), *Handbook of experimental economics results* (Vol. 1) (pp. 1061-1073). New York: Elsevier.
- Euteneuer, F., Schaefer, F., Stürmer, R., Boucsein, W., Timmermann, 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. *Neuropsychologica*, 47, 2882-2890.
- Fischer, L. & Wiswede, G. (2009). *Grundlagen der Sozialpsychologie* (3. neubearbeitete Aufl.). München: Oldenbourg.

- Frischenschlager, O., Mitterauer, L. & Haidinger, G. (2005). Leistungsfaktoren als potenzielle Auswahlkriterien im Medizinstudium. *Zeitschrift für Hochschuldidaktik*, 6, 34-41.
- Fonagy, P., Gergely, G., Jurist, E.L. & Target, M. (2008). *Affektregulierung, Mentalisierung und die Entwicklung des Selbst*. Stuttgart: Clett-Kotta.
- 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 stereoids: Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cogition*, 45, 392-414.
- Hänsgen, K.-D. & Spicher, B. (2007). *EMS Eignungstest für das Medizinstudium 2007 - Bericht 13 über die Durchführung und Ergebnisse 2007*. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.
- Harbaugh, W. T., Krause, K. & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, 5 (1), 53-84.
- Herkner, W. (2003). *Lehrbuch Sozialpsychologie*. Bern: Hans Huber.
- Herrmann, D. J., Crawford, M. & Holdworth, M. (1992). Gender-linked differences in everyday memory performance. *British Journal of Psychology*, 83, 221-231.
- Ionescu, M. (2000). Sex differences in memory estimates for pictures and words. *Psychological Reports*, 87, 315-322.
- Ionescu, M. (2002). Sex differences in memory estimates, Revisited. *Psychological Reports*, 91, 167-172.
- ITB Consulting (Hrsg.). (2008). Test für medizinische Studiengänge II: Originalversion II des TMS (5. Aufl.). Göttingen: Hogrefe.

- Johns, M., Schmader, T. & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, 16, 175-179.
- Jungermann, H., Pfister, H.-R. & Fischer, K. (2010). *Die Psychologie der Entscheidung. Eine Einführung* (3. Aufl.). Spektrum Akademischer Verlag: Heidelberg.
- Keller, J. (2008). Stereotype als Bedrohung. In L. E. Petersen & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorie, Befunde und Interventionen* (S. 88-96). Basel: Beltz Psychologie Verlags Union.
- Keller, J. & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Bulletin*, 29, 371-381.
- Kirchler, E. (1999). *Wirtschaftspsychologie* (2. überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Klauer, K. C. (2008). Soziale Kategorisierung und Stereotypisierung. In L. E. Petersen & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorie, Befunde und Interventionen* (S. 23-32). Basel: Beltz Psychologie Verlags Union.
- Kramer, J. H., Kaplan, E., Delis, D.C., O'Donnell L. & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology*, 11 (3), 577-584.
- Kubinger, K. D. (2006). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Lefrancois, G. R. (2006). *Psychologie des Lernens* (4. Aufl.). Heidelberg: Springer.
- Lerch, H.-J., Rauch, A. & Schlesier, A. (2000). *Risikobereitschaft und Entscheidungsunsicherheit. Eine Analyse von Zusammenhängen mit Alter, Geschlecht, IQ und Schulleistungen*. Regensburg: S. Roderer Verlag.

- Levin, I. P., Synder, M. A. & Chapman, D. P. (1988). The interaction of experiamental and situational factors and gender in a simulated risky decision-making task. *The Journal of Psychology*, 122, 173-181.
- Mallinger, R., Holzbaaur, Chr., Dierich, M., Heidegger, M., Zeller, J., Hänsngen, K.-D. & Spicher, B. (2009). *EMS Eignungstest für das Medizinstudium in Österreich. Medizinische Universität Innsbruck, Medizinische Universität Wien*. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.
- Martens, A., Johns, M., Greenberg, J. & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, 42, 236-243.
- 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.
- McGlone, M. S. & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, 27 (486-493).
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychological Review*, 63, 81-97.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appelton-Century-Crofts.
- Petersen, L. E. & Six, B. (Hrsg.). (2008). *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen*. Basel: Beltz Psychologie Verlags Union.
- Rost, R. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. überarbeitete und erweiterte Aufl.). Bern: Hans Huber

- Rydell, R. J., McConell, A.R. & Beilock, S.L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and Working Memory. *Journal of Personality and Social Psychology*, 96 (5), 949-966.
- Schermer, F. J. (2006). *Lernen und Gedächtnis* (4. überarbeitete und erweiterte Aufl.). Stuttgart: Kohlhammer.
- Schmader, T. & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85 (3), 440-452.
- Schuler, H. & Prochaska, M. (2001). *Leistungsmotivation (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.
- Seel, N. M. (2003). *Psychologie des Lernens* (2. aktualisierte und erweiterte Aufl.). Ernst Reinhardt: München.
- 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., Litzenberger, M. & Haiden, D. (2007). Bildungswissenschaftliche und psychologische Aspekte von Auswahlverfahren. In C. Badelt, W. Wegscheide & H. Wultz (Hrsg.), *Hochschulzugang in Österreich* (S. 479-552). Graz: Grazer Universitätsverlag.
- Spiel C., Schober, B. & Litzenberger, M. (2008). *Evaluation der Eignungstests für das Medizinstudium in Österreich*. Projektbericht für das Bundesministerium für Wissenschaft und Forschung. Wien.

- 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.
- Stone, J., Lynch, C.I., Sjomeling, M. & Darley, J.M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, 77, 1213-1227.
- Stumpf, H. & Fay, E. (1983). *Schlauchfiguren. Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens. Handanweisung.* Göttingen: Hogrefe.
- Trahan, D. E. & Quintana, J.W. (1990). Analysis of gender effects upon verbal and visual memory performance in adults. *Archives of Clinical Neuropsychology*, 5, 325-334.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp.381-403). New York: Academic Press.
- Von Klebelsberg, D. (1969). *Risikoverhalten als Persönlichkeitsmerkmal.* Bern: Hans Huber.
- West, R. L., Crook, T.H. & Barron, K.L. (1992). Everyday memory performance across the life span: Effects of age and noncognitive individual differences. *Psychology and Aging*, 7, 72-82.
- Wiling, U., Schmöger, M., Müller, C. & Auff. E. Theory of mind stories. *In Vorbereitung.*
- Zentrum für Testentwicklung und Diagnostik (ZTD). (2009). *TestInfo '09. Eignungstest für das Medizinstudium (EMS). Version A.* [Broschüre]. Innsbruck, Wien: Medizinische Universität.
- Zimbardo, P. G. & Gerrig, R.J. (2004). *Psychologie* (16. aktualisierte Aufl.). München: Pearson Studium.

8 ABBILDUNGSVERZEICHNIS

<i>Abbildung 1: Game of Dice Task (GDT). Nach der Entscheidung für eine Zahlenkombination wird gewürfelt und angezeigt, ob der Wurf erfolgreich oder nicht erfolgreich und der Gewinn/ Verlust ist am Bildschirm ersichtlich (Brand et al., 2005).....</i>	<i>51</i>
<i>Abbildung 2: Beispielaufgabe Wissenstest Figuren Lernen.....</i>	<i>59</i>
<i>Abbildung 3: Beispielaufgabe Wissenstest Fakten Lernen.....</i>	<i>59</i>
<i>Abbildung 4: Beispielaufgabe EMS- ähnlicher Untertest Fakten Lernen Einpräge- und Reproduktionsphase (ZTD, 2009).....</i>	<i>61</i>
<i>Abbildung 5: Beispielaufgabe EMS- ähnlicher Untertest Figuren Lernen Einpräge- und Reproduktionsphase (ZTD, 2009).....</i>	<i>61</i>
<i>Abbildung 6: Beispielaufgabe EMS- ähnlicher Untertest Quantitative und Formale Probleme (ZTD, 2009).....</i>	<i>62</i>
<i>Abbildung 7: Beispielaufgabe EMS- ähnlicher Untertest Schlauchfiguren (ZTD, 2009).....</i>	<i>62</i>
<i>Abbildung 8: Beispielaufgabe EMS- ähnlicher Untertest Medizinisch- naturwissenschaftliches Grundverständnis (ZTD, 2009).....</i>	<i>63</i>
<i>Abbildung 9: Beispielaufgabe EMS- ähnlicher Untertest Tabellen und Diagramme (ZTD, 2009) ...</i>	<i>64</i>
<i>Abbildung 10: Beispielaufgaben Leistungsmotivation (Schuler & Prohaska, 2001).....</i>	<i>65</i>
<i>Abbildung 11: Beispielaufgaben Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999) ..</i>	<i>66</i>
<i>Abbildung 12: Game of Dice Task (Brand et al., 2005).....</i>	<i>66</i>
<i>Abbildung 13: Anteilswerte der Geschlechter in Prozent.....</i>	<i>69</i>
<i>Abbildung 14: Anteilswerte der Nationalität in Prozent.....</i>	<i>70</i>
<i>Abbildung 15: Lebensalter der UntersuchungsteilnehmerInnen (1086 gültige Fälle).....</i>	<i>72</i>
<i>Abbildung 16: Verteilung Wiederholen in Abhängigkeit vom Geschlecht</i>	<i>77</i>
<i>Abbildung 17 Wunsch auf ein Medizinstudium in Abhängigkeit vom Geschlecht</i>	<i>78</i>
<i>Abbildung 18: Beschäftigung mit dem EMS und Geschlecht</i>	<i>81</i>
<i>Abbildung 19: Beschäftigung mit EMS abhängig vom Schulzweig</i>	<i>83</i>
<i>Abbildung 20: Verteilung der TeilnehmerInnen in die Stereotype Threat Bedingungen (Aktivierung, Neutral, Aufklärung und Negativaktivierung) in Abhängigkeit vom Geschlecht</i>	<i>86</i>
<i>Abbildung 21: Zuweisung der TeilnehmerInnen in die Risikobedingungen (Raten, Ohne Instruktion und Sicher) in Abhängigkeit vom Geschlecht.....</i>	<i>88</i>
<i>Abbildung 22: Stereotype Threat und Risikoinstruktion</i>	<i>89</i>
<i>Abbildung 23: Wissenstest Aufgabe Diagramme und Tabellen.....</i>	<i>101</i>
<i>Abbildung 24: Wechselwirkungsdiagramm- Interaktion zwischen Geschlecht und Risikoverhalten</i>	<i>113</i>
<i>Abbildung 25: Stereotype Threat Bedingungen und Untertest Fakten Lernen</i>	<i>117</i>

Abbildung 26: Untertest Figuren Lernen in Abhängigkeit Medizinstudium	121
Abbildung 27: NS-GDT Score in Abhängigkeit vom Geschlecht und Risikoinstruktion	127

9 TABELLENVERZEICHNIS

Tabelle 1: Untertests im EMS-Aufnahmeverfahren (Hängsen & Spicher, 2007)	13
Tabelle 2: Studiendesign und Zuweisung in die Versuchsbedingungen	55
Tabelle 3: Verteilung der Geschlechter über die Gesamtstichprobe	69
Tabelle 4: Chi- Quadrat Test Geschlechterverteilung über die Gesamtstichprobe	69
Tabelle 5: Häufigkeiten für Lebensalter über die Gesamtstichprobe	70
Tabelle 6: Chi- Quadrat Test für das Lebensalter	71
Tabelle 7: Häufigkeiten für besuchte Schulzweige	72
Tabelle 8: Chi- Quadrat Test für besuchte Schulzweige	73
Tabelle 9: Häufigkeitsverteilung und Anteilswerte der besuchten Schulzweige	74
Tabelle 10: Häufigkeitstabelle Wiederholen einer Schulstufe	74
Tabelle 11: Chi- Quadrat Test für Wiederholen einer Schulstufe	75
Tabelle 12: Kontingenztafel Wiederholen einer Schulklasse in Abhängigkeit vom Geschlecht	76
Tabelle 13: Häufigkeiten für den Wunsch auf ein Medizinstudium	77
Tabelle 14: Chi- Quadrat Test für den Wunsch auf ein Medizinstudium	78
Tabelle 15 Wunsch Medizin zu studieren in Abhängigkeit vom Geschlecht	79
Tabelle 16 Vierfeldertafel Wunsch Medizin zu studieren und EMS- Beschäftigung	80
Tabelle 17: Vierfeldertafel Beschäftigung mit EMS und Geschlecht	82
Tabelle 18: Vierfeldertafel Beschäftigung mit EMS und Schulzweig (naturwissenschaftlich)	84
Tabelle 19: Häufigkeiten Stereotype Threat Bedingungen	85
Tabelle 20: Chi- Quadrat Test für Stereotype Threat Bedingungen	85
Tabelle 21: Kontingenztafel Stereotype Threat und Geschlecht	86
Tabelle 22: Häufigkeiten Risikobedingungen	87
Tabelle 23: Chi- Quadrat Test für Risikobedingungen	87
Tabelle 24: Kontingenztafel Risikoverhalten und Geschlecht	88
Tabelle 25: Kontingenztafel Stereotype Threat und Risikoverhalten	90
Tabelle 26: Statistische Kennwerte und Koeffizienten der EMS- ähnlichen Untertests	91
Tabelle 27: Lösungswahrscheinlichkeiten im Untertest Figuren Lernen	92
Tabelle 28: Lösungswahrscheinlichkeiten im Untertest Fakten Lernen	92
Tabelle 29: Kontingenztafel Stereotype Threat (4) und Wissenstest Aufgabe Fakten Lernen	97

<i>Tabelle 30: Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe Figuren Lernen</i>	<i>98</i>
<i>Tabelle 31: Kennwerte und Prüfgrößen der Selbsteinschätzung im Wissenstest in Abhängigkeit vom Geschlecht.....</i>	<i>100</i>
<i>Tabelle 32: Vergleich Cronbach's Alpha Reliabilitäten der EMS Untertests (Hänsgen & Spicher, 2007) und Game of Dice Task (GDT)</i>	<i>102</i>
<i>Tabelle 33: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Figuren Lernen</i>	<i>104</i>
<i>Tabelle 34: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Fakten Lernen</i>	<i>105</i>
<i>Tabelle 35: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Quantitative und Formale Probleme.....</i>	<i>105</i>
<i>Tabelle 36: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Medizinisch-naturwissenschaftliches Grundverständnis</i>	<i>106</i>
<i>Tabelle 37: Reliabilitäten der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Untertest Diagramme und Tabellen.....</i>	<i>106</i>
<i>Tabelle 38: Reliabilitäten der Faktorstufenkombination (Stichprobenumfang, Itemanzahl) EMS-ähnlicher Unterst Schlauchfiguren.....</i>	<i>107</i>
<i>Tabelle 39: Reliabilität mittels Cronbach's Alpha EMS- ähnlicher Gesamttest</i>	<i>107</i>
<i>Tabelle 40: Reliabilitäten mittels Cronbach's Alpha EMS- ähnlicher Gesamttest</i>	<i>108</i>
<i>Tabelle 41: Interkorrelationsmatrix der Produkt- Moment Korrelation r der EMS- ähnlichen Untertests (N = 1093).....</i>	<i>108</i>
<i>Tabelle 42: Kennwerte EMS- ähnlicher Untertest Figuren Lernen</i>	<i>111</i>
<i>Tabelle 43: Test der Zwischensubjekteffekte, abhängige Variable: EMS Figuren Lernen</i>	<i>111</i>
<i>Tabelle 44: Deskriptiv Statistik * EMS- ähnlicher Untertest Fakten Lernen</i>	<i>115</i>
<i>Tabelle 45: Tests der Zwischensubjekteffekte, abhängige Variable: EMS Fakten</i>	<i>116</i>
<i>Tabelle 46: Vierfelder- Tafel Wunsch Medizin zu studieren und spezifische Vorbereitung auf den EMS</i>	<i>119</i>
<i>Tabelle 47: Test der Zwischensubjekteffekte Untertest Figuren Lernen in Abhängigkeit vom Wunsch Medizin zu studieren und spezifischer Vorbereitung auf den EMS</i>	<i>120</i>
<i>Tabelle 48: Deskriptiv Statistik Wunsch, Medizin zu studieren und spezifische Vorberitung auf den EMS im Unterest Figuren Lernen</i>	<i>121</i>
<i>Tabelle 49: Test der Zwischensubjekteffekte Untertest Fakten Lernen in Abhängigkeit vom Wunsch Medizin zu studieren und spezifischer Vorbereitung auf den EMS</i>	<i>122</i>
<i>Tabelle 50: Deskriptiv Statistik Wunsch, Medizin zu studieren und spezifische Vorbereitung auf den EMS im Untertest Fakten Lernen</i>	<i>123</i>
<i>Tabelle 51: Kennwerte Bedingungen Risikoinstruktion.....</i>	<i>125</i>

<i>Tabelle 52: Test der Zwischensubjekteffekte, abhängige Variable: NS-GDT</i>	<i>126</i>
--	------------

10 ANHANG

Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe Naturwissenschaftliches Grundverständnis

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

Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe Schlauchfiguren

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

**Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe
Quantitative und Formale Probleme**

Kreuztabelle Stereotype * Quantitative und Formale Probleme						
			Formal			Gesamt
			nein	ja	entgegen	
Stereo	Aktivierung	Anzahl	60	207	22	289
		erwartet	90,7	157,9	40,4	289,0
		% innerhalb von	20,8%	71,6%	7,6%	100,0%
		Stand. Residuen	-3,2	3,9	-2,9	
	Neutral	Anzahl	86	142	26	254
		erwartet	79,7	138,8	35,5	254,0
		% innerhalb von	33,9%	55,9%	10,2%	100,0%
		Stand. Residuen	,7	,3	-1,6	
	Aufklärung	Anzahl	141	114	22	277
		erwartet	86,9	151,3	38,7	277,0
		% innerhalb von	50,9%	41,2%	7,9%	100,0%
		Stand. Residuen	5,8	-3,0	-2,7	
	Negativaktivierung	Anzahl	52	127	81	260
		erwartet	81,6	142,0	36,4	260,0
		% innerhalb von	20,0%	48,8%	31,2%	100,0%
		Standardisierte Residuen	-3,3	-1,3	7,4	
Gesamt		Anzahl	339	590	151	1080
		erwartet	339,0	590,0	151,0	1080,0
		% innerhalb von	31,4%	54,6%	14,0%	100,0%

Kontingenztafel Stereotype Threat Bedingungen (4) und Wissenstest Aufgabe
Diagramme und Tabellen

Stereotype * Diagramme und Tabellen Kreuztabelle						
			Tabellen			Gesamt
			nein	ja	entgegen	
Stereo	Aktivierung	Anzahl	52	223	14	289
		Erwartet	85,6	170,5	32,9	289,0
		% innerhalb von	18,0%	77,2%	4,8%	100,0%
		Stand. Residuen	-3,6	4,0	-3,3	
	Neutral	Anzahl	95	139	20	254
		Erwartet	75,3	149,8	28,9	254,0
		% innerhalb von	37,4%	54,7%	7,9%	100,0%
		Stand. Residuen	2,3	-,9	-1,7	
	Aufklärung	Anzahl	125	123	29	277
		Erwartet	82,1	163,4	31,5	277,0
		% innerhalb von	45,1%	44,4%	10,5%	100,0%
		Stand. Residuen	4,7	-3,2	-,5	
	Negativaktivierung	Anzahl	48	152	60	260
		Erwartet	77,0	153,4	29,6	260,0
		% innerhalb von	18,5%	58,5%	23,1%	100,0%
		Stand. Residuen	-3,3	-,1	5,6	
Gesamt		Anzahl	320	637	123	1080
		Erwartet	320,0	637,0	123,0	1080,0
		% innerhalb von	29,6%	59,0%	11,4%	100,0%

Statistische Kennwerte Fremdeinschätzung im Wissenstest

	Stereotype											
	A Aktivierung			B neutral			C Aufklärung			D Negativaktivierung		
	Gültig g	M	SD	Gültig	M	SD	Gültig	M	SD	Gültig	M	SD
FigM	289	42.78	24.143	255	49.47	8.000	279	50.06	8.606	261	55.95	25.291
FigF	289	57.25	24.122	255	50.76	7.981	279	49.90	8.730	261	44.09	25.308
FaktM	290	33.69	23.036	254	48.13	13.263	278	46.98	14.125	260	61.13	26.790
FaktF	290	66.22	23.196	254	51.87	13.263	278	53.00	14.221	260	38.60	26.827
NatM	290	59.21	22.244	254	52.78	9.611	277	52.12	9.706	259	41.60	24.953
NatF	290	40.76	22.257	254	47.26	9.585	277	47.88	9.668	260	58.22	25.085
DiagM	289	69.71	19.777	254	54.57	9.236	277	55.09	11.511	260	36.09	25.020
DiagF	289	30.35	19.829	254	45.44	9.238	277	44.84	11.142	260	64.12	24.786
QuaM	289	68.56	21.207	254	53.70	12.333	277	56.45	11.803	260	42.84	26.155
QuaF	289	31.54	21.322	254	46.10	11.863	277	43.58	11.810	260	57.06	26.378
SchlM	289	71.65	20.886	254	53.08	12.247	277	54.62	12.270	260	36.35	26.767
SchlF	289	28.45	21.023	254	47.00	12.137	277	45.36	12.274	260	63.89	26.499
ToMM	288	21.71	17.165	254	34.25	14.394	276	31.68	13.803	257	22.19	18.771
ToMF	288	78.36	17.078	254	65.99	14.404	276	68.32	13.803	257	77.82	18.744

Curriculum Vitae

HELENE ANTONIA LAGGER

Mariannengasse 17/ 7

1090 Wien

Persönliche Daten:

Geboren am 29.11.1981 in Klagenfurt

Österreichische Staatsbürgerin, ledig

Berufliche Ausbildung & Schulbildung:

Psychologiestudium an der Universität Wien

seit Wintersemester 01/02

Oberstufenrealgynasium der Ursulinen von 1996-2000

Berufliche Tätigkeit:

x3 projects – PR Agentur und VeranstaltungsGmbH.

Assistentin

laufend seit 2008

Praktika:

Mitarbeit auf der **neuroonkologischen Abteilung des Akh** im Rahmen des
„Clinic Nanny“ Projektes

WS 2008- SS 2010

Praktikum bei **e.motion, Egotherapie, Otto- Wagner Spital** Wien, Steinbachhof
Juni 2009- August 2009- Ausmaß 240 Stunden