



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Zum Einfluss von Geschlecht, Stereotype Threat und
Risikoverhalten auf die Leistungen bei EMS-
ähnlichen Untertests und die Berücksichtigung der
Selbstwirksamkeitserwartung

verfasst von

Anita Teufl

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 298

Studienrichtung lt. Studienblatt:

Diplomstudium Psychologie

Betreuerin:

Ao. Univ. Prof. Dr. Mag. Ulrike Willinger

Danksagung

Als erstes möchte ich mich bei Frau Ao. Univ. Prof. Dr. Mag. Ulrike Willinger für die fachkundige Betreuung meiner Diplomarbeit, die freundliche Unterstützung und die vielen konstruktiven wissenschaftlichen Impulse bedanken.

Mein größter Dank gebührt meinen Eltern Maria & Leopold Teufl, die mich seit jeher in jeglicher Weise unterstützt, bekräftigt und gefördert haben. Ihr habt mir soviel auf den Weg mitgegeben und ich danke Euch von Herzen für Eure Liebe, Kraft und Unterstützung.

Besonderer Dank gilt auch meinem Bruder Ing. Reinhard Teufl für seine immer währende Hilfsbereitschaft nicht nur in computertechnischen Angelegenheiten, seine Zuversicht und Rückhalt. Meinen Großeltern Maria & Josef, die immer für mich da waren und mich unterstützt haben, danke ich von Herzen. Ein Dankeschön ebenso an meine restliche Familie, insbesondere an meine Cousine Carina für ihr stets offenes Ohr und Frohsinn.

An dieser Stelle möchte ich mich bei meinen FreundInnen ganz herzlich für ihre Freundschaft, Unterstützung und die lustigen, ablenkenden Abende in Wien, NÖ sowie fernab bedanken. Meiner besten Freundin Ingrid danke ich insbesondere für ihren Zuspruch, ihre moralische Unterstützung und Motivation.

Großer Dank gebührt meinen Diplomarbeitkolleginnen Silvia, Katrin, Helene und Gabi für die gegenseitige Aufmunterung sowie die angenehme und konstruktive Zusammenarbeit während der Durchführung dieser wissenschaftlichen Untersuchung.

Ein großes Dankeschön gilt den engagierten Lehrkräften, bei denen unser Projekt Anklang fand und insbesondere den 1093 AHS-SchülerInnen, die teilweise beim Testzeitpunkt sehr kurz vor der stressreichen Zeit der Matura standen und trotzdem an der Untersuchung bereitwillig teilgenommen haben.

Abschließend bedanke ich mich bei all denjenigen, die mich auf diesem Weg begleitet und inspiriert haben.

Inhaltsverzeichnis

1. Einleitung	1
2. Eignungstest für das Medizinstudium	3
3. Stereotype Threat	8
3.1 Einführung	8
3.2 Begriffsbestimmung	9
3.3 Auswirkungen auf die Leistung	9
3.4 Vermittelnde Mechanismen.....	13
3.5 Leistungsminderungen reduzieren	15
4. Selbstwirksamkeitserwartung	18
4.1 Begriffsbestimmung	18
4.2 Struktur der Selbstwirksamkeitserwartungen	19
4.3 Quellen der Selbstwirksamkeitserwartungen.....	21
4.4 Selbstwirksamkeitserwartung und Leistung	23
4.5 Selbstwirksamkeitserwartung und Stereotype Threat.....	28
5. Risikoverhalten	30
6. Zielsetzungen, Fragestellungen und Hypothesen	34
6.1 Zielsetzungen und Fragestellungen	34
6.2 Hypothesen.....	34
7. Methode	37
7.1 Untersuchungsplan und intendierte Stichprobe.....	37
7.2 Erhebungsinstrumente	40
7.2.1 Soziodemographischer Fragebogen	40
7.2.2 Wissenstest	41
7.2.3 EMS-ähnlicher Test.....	42
7.2.4 Theory of Mind - Aufgaben.....	48
7.2.5 Leistungsmotivationsinventar	48
7.2.6 Skala zur Allgemeinen Selbstwirksamkeitserwartung.....	49
7.2.7 Game of Dice Task	50
8. Untersuchung	52
8.1 Untersuchungsdurchführung.....	52
8.2 Auswertungsverfahren	53

8.3	Stichprobenbeschreibung	56
8.3.1	Geschlecht.....	56
8.3.2	Alter	57
8.3.3	Schulstandort und Staatsbürgerschaft.....	58
8.3.4	Ausbildungszweig	59
8.3.5	Wiederholung einer Schulstufe.....	61
8.3.6	Ausbildung der Eltern.....	62
8.3.7	Intention, Medizin zu studieren.....	66
8.3.8	Befassung mit dem EMS.....	67
8.3.9	Verteilung der UntersuchungsteilnehmerInnen in den Untersuchungsbedingungen	69
9.	Ergebnisse	74
9.1	Kennwerte der Erhebungsinstrumente	74
9.1.1	EMS-ähnliche Untertests.....	74
9.1.2	Lösungswahrscheinlichkeiten und Itemtrennschärfen in den EMS- ähnlichen Subtests	78
9.1.3	Selbstwirksamkeitserwartung.....	86
9.1.4	Game of Dice Task	88
9.2	Überprüfung der Wirksamkeit der Stereotype Threat- Untersuchungsbedingungen.....	89
9.3	Hypothesenbezogene Ergebnisse	101
9.3.1	Selbstwirksamkeitserwartung.....	101
9.3.2	Leistungen in EMS-ähnlichen Untertests.....	105
9.3.3	Leistungen in EMS-ähnlichen Untertests unter Berücksichtigung der SWE als Kovariate.....	114
9.3.4	Risikoverhaltens-Tendenz.....	120
10.	Diskussion	123
11.	Zusammenfassung	131
12.	Abstract.....	134
13.	Literaturverzeichnis.....	136
14.	Abbildungsverzeichnis.....	144
15.	Tabellenverzeichnis.....	146
16.	Anhang.....	151

Anmerkung

Die wissenschaftliche Untersuchung der vorliegenden Diplomarbeit basiert auf einer umfassenden Studie, an deren Durchführung unter der Leitung von Frau Ao. Univ. Prof. Dr. Mag. Ulrike Willinger, fünf Diplomandinnen der Fakultät für Psychologie involviert waren. Die Arbeiten meiner Kolleginnen (Katrín Anzirk, Silvia Hameseder, Gabriele Hangl und Helene Lagger) und mir fußen somit auf demselben Datensatz, wobei der Schwerpunkt jeder Diplomandin auf einem jeweils anderen Untersuchungsgegenstand liegt.

1. Einleitung

Im Rahmen des *Eignungstests für das Medizinstudium* (EMS) können größere Leistungsunterschiede zwischen den weiblichen und männlichen StudienbewerberInnen aus Österreich aufgezeigt werden, die zuungunsten der Frauen ausfallen (Mallinger et al., 2007, 2010).

In der vorliegenden Arbeit werden Faktoren untersucht, die möglicherweise einen Einfluss auf die Leistungen bei EMS-ähnlichen Untertests ausüben könnten. Neben dem Geschlecht nimmt das Phänomen *Stereotype Threat* einen weiteren Stellenwert ein, der im Rahmen der Untersuchung analysiert wird.

Hierbei handelt es sich um die Bedrohung, dass man ein negatives Stereotyp sich selbst und anderen bestätigen könnte (Steele & Aronson, 1995). Frauen sind beispielsweise von negativen Stereotypen hinsichtlich ihrer Leistungsfähigkeit in Mathematik und den Naturwissenschaften betroffen (Spencer, Steele & Quinn, 1999; Steele, 1997). Untersuchungen verdeutlichen, dass das Phänomen *Stereotype Threat* die Testleistung von Frauen (u.a. in Mathematik) beeinträchtigen kann (z.B.: Smith & White, 2002).

Ein dritter Faktor, der in der Untersuchung analysiert wird, kennzeichnet die Risikoverhaltens-Tendenz: Geschlechtsspezifische Unterschiede im *Risikoverhalten* (z.B. Raten), könnten ein etwaiger Erklärungsansatz für geringere Leistungen von Frauen bei Multiple Choice-Tests darstellen (z.B.: Ben-Shakhar & Sinai, 1991; Bolger & Kellaghan, 1990).

Darüber hinaus steht in der vorliegenden Arbeit das Konstrukt der *Selbstwirksamkeitserwartung* im Fokus. Viele Untersuchungen belegen einen positiven Zusammenhang zwischen der Selbstwirksamkeitserwartung und Leistung (z.B.: Schwarzer, Mueller & Greenglass, 1999) auf. Des Weiteren stellt die Selbstwirksamkeitserwartung einen starken Prädiktor der Leistung (z.B.: Mathematik) dar (z.B.: Pajares & Kranzler, 1995), weist einen positiven Zusammenhang mit dem Verbleib in intendierten Studiengängen auf (Lent, Brown & Larkin, 1984) und kann als Prädiktor der wahrgenommenen beruflichen Karrieremöglichkeiten angenommen werden (Lent, Brown & Larkin, 1986).

Im Rahmen der vorliegenden Arbeit wird analysiert, ob das Geschlecht, Stereotype Threat und Risikoverhalten die Leistungen in EMS-ähnlichen Untertests beeinflussen. Weiters erfolgt die Überprüfung des Zusammenhangs zwischen der Selbstwirksamkeitserwartung und den Leistungen in EMS-ähnlichen Untertests sowie möglicher geschlechtsspezifische Unterschiede in der Selbstwirksamkeitserwartung. Darüber hinaus wird untersucht, ob ein etwaiger Einfluss der Selbstwirksamkeitserwartung als Kovariate auf die Leistungen in den EMS-ähnlichen Untertests in Abhängigkeit der drei genannten Faktoren angenommen werden kann.

Der nachfolgende *theoretische Teil* der Arbeit widmet sich zunächst dem Eignungstest für das Medizinstudium. Daran anschließend erfolgt die Darstellung des Phänomens Stereotype Threat. In diesem Kapitel werden die negativen Auswirkungen des Phänomens auf die Leistung von Frauen thematisiert und etwaige vermittelnde Mechanismen des Phänomens sowie mögliche Strategien zur Reduzierung aufgezeigt. Als nächstes wird das Konstrukt der Selbstwirksamkeitserwartung erläutert. Neben der Darlegung der Struktur und Quellen der Selbstwirksamkeitserwartungen wird der Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Leistung beleuchtet. Abschließend erfolgt die Auseinandersetzung hinsichtlich einer etwaigen vermittelnden Rolle der Selbstwirksamkeitserwartung auf die Stereotype Threat-Effekte und die Analyse geschlechtsspezifischer Unterschiede im Risikoverhalten.

Im *empirischen Teil* der Arbeit werden zuerst Untersuchungsplan und intendierte Stichprobe dargelegt. Anschließend erfolgen die Darstellung der eingesetzten Erhebungsinstrumente, der Durchführung der Untersuchung sowie der statistischen Auswertungsverfahren. Weiters wird die Stichprobe hinsichtlich interessierender sozialstatistischer Merkmale beschrieben sowie die Untersuchungsergebnisse dargestellt, diskutiert und zusammengefasst.

Theoretischer Teil

2. Eignungstest für das Medizinstudium

An den Medizinischen Universitäten in Wien und Innsbruck erfolgt seit dem Studienjahr 2006 gemäß der „Novelle zum Universitätsgesetz 2002, gemeinsam eine kapazitätsorientierte Studienplatzvergabe für ALLE StudienbewerberInnen“ (Mallinger et al., 2010, S. 4).

Ausgangspunkt hierfür stellte das Urteil des Europäischen Gerichtshofs (7. Juli 2005) zum österreichischen Hochschulzugang dar, das ein übermäßiges Zuströmen ausländischer Studierender nach sich zog. Den BewerberInnen stehen an der Medizinischen Universität in Wien 660 Studienplätze für Humanmedizin sowie 80 für Zahnmedizin und an der Medizinischen Universität in Innsbruck 360 Studienplätze für Humanmedizin sowie 40 für Zahnmedizin zur Verfügung (Mallinger et al., 2010).

Das Fach der Medizin wird mit einer großen Zukunftssicherheit verbunden und stellt ein sehr gefragtes Studium dar (Hänsgen, 2007). Für das Studienjahr 2010/11 gab es an den Medizinischen Universitäten in Wien und Innsbruck insgesamt 5967 BewerberInnen (44,4 % Männer und 55,6% Frauen): 3294 aus Österreich und 2673 aus dem Ausland (Mallinger et al., 2010). Zum Schutz der Homogenität des österreichischen Bildungssystems wurde gemäß § 124b Abs. 5 Universitätsgesetz 2002 festgelegt, dass 75 Prozent der Gesamtstudienplätze für InhaberInnen in Österreich ausgestellter Reifezeugnisse und 20 Prozent für EU-BürgerInnen bestimmt sind (Bundesministerium für Wissenschaft und Forschung [BMWF], 2002). Die Ermittlung derjenigen BewerberInnen, die einen Studienplatz erhalten, erfolgt seit 2006 an den Medizinischen Universitäten in Wien und Innsbruck anhand eines Eignungstests (Mallinger et al., 2010). Das verwendete Verfahren, der *Eignungstest für das Medizinstudium* (EMS), stellt eine Adaptation des Tests für Medizinische Studiengänge (TMS) dar, dessen Teststruktur in Deutschland wissenschaftlich erprobt wurde und in der Schweiz bereits seit 1998 den BewerberInnen für ein Medizinstudium vorgegeben wird (Hänsgen & Spicher, 2000, 2010).

Der EMS bestand bis einschließlich 2004 aus insgesamt neun Untertests (Hänsgen & Spicher, 2010). Seit 2005 erfolgt auch die Erhebung der Planungskompetenz (Subtest:

Planen und Organisieren), entsprechend neuer Anforderungen des Medizinstudiums (Hänsgen, 2007; Hänsgen & Spicher, 2010). Im Rahmen von 10 Untertests mit insgesamt 198 Aufgaben (siehe Tabelle 1), überprüft der EMS unterschiedliche studienrelevante Fähigkeiten (Hänsgen & Spicher, 2010).

Tabelle 1: EMS-Teststruktur (nach Hänsgen & Spicher, 2010, S. 41)

Bezeichnung der Untertests	Aufgabenanzahl	Studienrelevante Fähigkeiten (Hänsgen & Spicher, 2010, S. 93ff.)	Bearbeitungszeit (in Minuten)
1) Quantitative und formale Probleme	20	Quantitative und formale Kompetenzen bei medizinisch-naturw. Fragestellungen	50
2) Schlauchfiguren	20	räumliches Vorstellungsvermögen	12
3) Textverständnis	18	Aufnahme und Verarbeitung komplexer Texte	45
4) Planen und Organisieren	20	Planungs- und Organisationsfähigkeiten	60
5) Konzentriertes und sorgfältiges Arbeiten	Blatt mit 1600 Zeichen	schnelles, sorgfältiges Arbeiten und Konzentration	8
	<i>Pause</i>		60
<i>Lernphase</i> Figuren lernen Fakten lernen	Vorlagen zum Einprägen		4 6
6) Medizinisch-naturwissenschaftliches Grundverständnis	20	Verständnis- und Schlussfolgerungsfähigkeit bei medizinisch-naturw. Fragestellungen	50
<i>Reproduzieren</i> 7) Figuren lernen	20	Merkfähigkeit von: - bestimmten Einzelheiten	5
8) Fakten lernen	20	- Fakten	7
9) Muster zuordnen	20	Musterausschnitte erkennen	18
10) Diagramme und Tabellen	20	Analyse- und Interpretationsfähigkeit bei Diagrammen u. Tabellen	50
Gesamt	198		315

Bei der Testung erhalten die BewerberInnen vormittags und nachmittags jeweils fünf Untertests zur Bearbeitung. Die Aufgaben der einzelnen Subtests sind aufsteigend nach Schwierigkeitsgrad geordnet und die Bearbeitung der Untertests findet zudem unter Zeitbeschränkung statt (siehe Tabelle 1), wobei die vorgegebene

Bearbeitungszeit in einzelnen Subtests relativ gering bemessen ist. Während von den zehn Untertests neun aus Multiple Choice Items mit vier oder fünf Antwortmöglichkeiten und immer nur einer korrekten Lösung bestehen, müssen im Subtest *Konzentriertes und sorgfältiges Arbeiten* auf einem Blatt mit 40 Zeichen pro Zeile (insgesamt 40 Zeilen) gemäß der in der Instruktion vorgegebenen Regel, Zeichen markiert werden (Zentrum für Testentwicklung und Diagnostik [ZTD], 2009).

Entscheidend für die erfolgreiche Bewältigung der Testaufgaben werden die Aneignung von Wissen sowie die Problemlösefähigkeit einer Person angesehen (Hänsgen & Spicher, 2010). Mittels des EMS wird kein Faktenwissen überprüft, hingegen soll eine Arbeitsprobe der Studierfähigkeit gegeben werden (Hänsgen, 2007). Die Studienplatzvergabe wird anhand des Testwertes, der sich aus der Addition der Punktwerte¹ in den zehn Subtests und der Standardisierung des Gesamtpunktwertes ergibt, durchgeführt (Hänsgen & Spicher, 2010). Im Rahmen des Einsatzes des EMS an den Medizinischen Universitäten in Wien und Innsbruck können jedoch beträchtliche Unterschiede in der Testleistung zwischen weiblichen und männlichen StudienbewerberInnen aus Österreich aufgezeigt werden: Frauen erreichen eine geringere Leistung als Männer (Mallinger et al., 2007, 2010).

Zudem wurde festgestellt, dass die BewerberInnen aus Österreich eine niedrigere Testleistung im Vergleich zu den BewerberInnen der EU (Quote) erzielen. Dieser Leistungsunterschied hat aufgrund der festgelegten Quote (siehe S. 3) keinen negativen Einfluss auf die Studienzulassung der österreichischen BewerberInnen. Ein anderes Bild zeichnet sich jedoch bei den Leistungsunterschieden zwischen den weiblichen und männlichen BewerberInnen aus Österreich hinsichtlich der Zulassung zum Medizinstudium ab (Mallinger et al., 2007).

Gegenüber der Bewerbungsquote erhalten weniger weibliche Bewerber einen Studienplatz als männliche. In Anbetracht der Ergebnisse des Testeinsatzes 2010 in Österreich zeigt sich bei den Untertests folgendes Bild: Männer erreichen bei sieben Untertests (*Quantitative und formale Probleme, Schlauchfiguren, Textverständnis, Planen und Organisieren, Konzentriertes und sorgfältiges Arbeiten, Medizinisch-naturwissenschaftliches Grundverständnis sowie Diagramme und Tabellen*) eine höhere Leistung als Frauen. Im Untertest *Muster zuordnen* sind keine geschlechtsspezifischen Leistungsunterschiede anzunehmen und bei den Untertests

¹ Eine richtig beantwortete Aufgabe wird mit einem Punkt und eine falsche mit null Punkten gewertet (ZTD, 2009).

Figuren lernen und *Fakten lernen* erzielen Frauen eine höhere Leistung im Vergleich zu Männern. Geschlechtsspezifische Leistungsunterschiede bei den Untertests können zwar auch in der Schweiz in einem vergleichbaren Muster beobachtet werden, allerdings in einem deutlich geringeren Ausmaß (Mallinger et al., 2010).

In einer Evaluationsstudie von Spiel, Schober und Litzenberger (2008) wurde der EMS im Rahmen des Auswahlverfahrens 2007/08 an den Medizinischen Universitäten in Wien und Innsbruck untersucht. Den insgesamt 3704 BewerberInnen für Humanmedizin wurde anschließend an den EMS ein zweiseitiger Fragebogen vorgelegt. Gemäß den teilnehmenden Personen sowie Ausschlusskriterien basiert die Untersuchung hauptsächlich auf einer Stichprobe von 3219 StudienbewerberInnen (43,3 % Männer und 56,7 % Frauen). Die Autorinnen stellten im Zuge der Überprüfung von neun EMS Subtests (ohne *Konzentriertes und sorgfältiges Arbeiten*) adäquate Itemschwierigkeiten fest, jedoch lagen die Trennschärfen von den insgesamt 178 analysierten Aufgaben bei 138 Items unter 0.30. Während für den Gesamttest eine Reliabilität von 0.92 ermittelt wurde, zeigen die neun Subtests keine hohen Reliabilitäten (≤ 0.75) auf. Die Ergebnisse der Fragebogenerhebung verdeutlichen, dass BewerberInnen, die zum Medizinstudium zugelassen wurden bessere Schulnoten sowohl in naturwissenschaftlichen Fächern als auch in Deutsch und Englisch aufweisen, als die nicht-erfolgreichen StudienbewerberInnen. Weiters kamen die Autorinnen zu dem Ergebnis, dass aufgenommene BewerberInnen eine höhere Zuversicht bekundeten, den EMS erfolgreich zu absolvieren, ein höheres Vertrauen in ihre mathematisch-naturwissenschaftlichen Fähigkeiten angaben sowie öfters an Aufnahmetests für ein Medizinstudium teilnahmen und sich in einem höheren Maß vorbereiteten im Vergleich zu den nicht zugelassenen BewerberInnen. Mehr als ein Drittel der StudienbewerberInnen gab an, den EMS als eher unfair einzuschätzen.

Die Berechtigung für den Einsatz des EMS als Zulassungskriterium zum Medizinstudium, liegt, gemäß Autoren, in der Prognosegüte des Studienerfolgs. Demnach sollen BewerberInnen aufgenommen werden, die das Medizinstudium sowohl erfolgreich als auch zügig abschließen (Hänsgen & Spicher, 2000, 2001).

Im Rahmen einer Evaluation des EMS in der Schweiz wurde der Zusammenhang zwischen dem erzielten EMS-Testwert (1266 StudentInnen des Studienjahres 1998 und 1999) und der Leistung bei der ersten sowie zum Teil auch der zweiten ärztlichen Vorprüfung überprüft. Die Ergebnisse zeigten einen positiven Zusammenhang zwischen dem EMS-Testwert und der erfolgreichen Absolvierung der ersten ärztlichen

Vorprüfung (erster Antritt) auf. Bezugnehmend auf die Geschlechtergruppen fand man, dass bei gleichen EMS-Testwerten sowohl für Männer als auch für Frauen dieselbe Bestehenswahrscheinlichkeit der ersten ärztlichen Vorprüfung angenommen werden kann. Geschlechtsspezifische Unterschiede konnten aber hinsichtlich der Bestehenswahrscheinlichkeit dieser ersten Vorprüfung festgestellt werden: Die Wahrscheinlichkeit, diese Prüfung erfolgreich zu absolvieren, liegt für Männer um sieben Prozent höher als für Frauen. Des Weiteren ergab die Evaluation, dass unabhängig vom Geschlecht, StudentInnen mit höheren EMS-Testwerten bessere Prüfungsnoten in der ersten ärztlichen Vorprüfung erzielen. Die Ergebnisse der zweiten ärztlichen Vorprüfung legen dar, dass StudentInnen, die diese Prüfung erfolgreich absolvierten, beim EMS eine höhere Leistung aufweisen im Vergleich zu StudentInnen, die die zweite Vorprüfung nicht erfolgreich absolvierten. Gemäß den Evaluationsergebnissen gelingt die Erfassung der Studieneignung mittels EMS (Hänsgen & Spicher, 2001).

Laut Mallinger et al. (2008) bestehe in Österreich bei Fortbestand ausgeprägter geschlechtsspezifischer Unterschiede in der EMS-Testleistung der österreichischen BewerberInnen, die (womöglich) aus bildungssoziologischen Benachteiligungen der weiblichen Bewerberinnen hervorgehen, Handlungsbedarf. In diesem Sinne würde eine für weibliche und männliche BewerberInnen getrennte Standardisierung des Punktwertes einen möglichen Ansatz darstellen, der jedoch mit dem Nachteil verbunden ist, dass gleiche EMS-Testleistungen für weibliche und männliche BewerberInnen nicht mit denselben Zulassungschancen einhergehen.

Im Rahmen des EMS-Testeinsatzes 2012 an der Medizinischen Universität Wien wurde dieser Ansatz durchgeführt (Medizinische Universität Wien [MedUni Wien], 2012a). An dieser Stelle ist darauf hinzuweisen, dass ab dem Studienjahr 2013/14 an den Medizinischen Universitäten in Wien, Innsbruck und Graz gemeinsam neue Aufnahmeverfahren angewendet werden (MedUni Wien, 2012b).

3. Stereotype Threat

3.1 Einführung

Unter dem Begriff *Stereotyp* versteht man nach Fiedler und Bless (2002): „Sozial geteilte Meinungen über Persönlichkeitsmerkmale und Verhaltensweisen von Mitgliedern einer sozialen Kategorie. Durch die Bildung von Stereotypen lässt man Individualität außer Acht“ (S. 134).

Im Hinblick auf den Leistungsbereich besteht beispielsweise bei Frauen das negative Stereotyp, wonach sie über geringere mathematische Fähigkeiten verfügen würden im Vergleich zu Männern. Neben Mathematik sind Frauen hauptsächlich auch in den Naturwissenschaften von negativen Stereotypen hinsichtlich ihrer Leistungsfähigkeit betroffen. In diesem Zusammenhang haben in Amerika zudem AfroamerikanerInnen mit negativen Stereotypen, die sich ihrerseits sowohl auf die meisten schulischen als auch akademischen Domänen erstrecken, zu kämpfen (Spencer, Steele & Quinn, 1999; Steele, 1997).

Steele (1997) weist darauf hin, dass Frauen in Mathematik, den Naturwissenschaften und in technischen Bereichen durchschnittlich geringere Leistungen im Vergleich zu Männern erreichen. In Bezug auf die Mathematikleistung verdeutlicht eine Metaanalyse, dass zu Beginn (Grundschule, Mittelschule) noch keine geschlechtsspezifischen Leistungsunterschiede beobachtet werden, jedoch ab etwa der Sekundarstufe steigen die Leistungsunterschiede zugunsten der Männer kontinuierlich bis ins Erwachsenenalter an (Hyde, Fennema & Lamon, 1990, zitiert nach Steele, 1997, S. 615). Beträchtliche Leistungsunterschiede zeigen sich auch zwischen afroamerikanischen und weißen SchülerInnen sowie StudentInnen (Steele, 1992, zitiert nach Steele & Aronson, 1995, S. 798).

In diesem Zusammenhang wurde der Frage nachgegangen, wie die Leistungsunterschiede erklärt werden können und es wurden verschiedenste Faktoren (z.B.: sozioökonomische) vermutet; jedoch konnten anhand dieser die Unterschiede nicht ausreichend geklärt werden (Steele, 1997; Steele & Aronson, 1995). In weiterer Folge wurde ein offenbar bedeutsamer Faktor in die Forschung eingebracht: „Stereotype Threat“ (Steele & Aronson, 1995).

3.2 Begriffsbestimmung

Mit dem Begriff *Stereotype Threat* (Bedrohung durch Stereotype) bezeichnet man die Bedrohung, dass man ein negatives Stereotyp (die eigene Gruppe betreffend) sich selbst und anderen, bestätigen könnte (Steele & Aronson, 1995).

Stereotype Threat kennzeichnet gemäß Steele (1997, S. 614) eine wahrgenommene Bedrohung:

[...] that arises when one is in a situation or doing something for which a negative stereotype about one's group applies. This predicament threatens one with being negatively stereotyped, with being judged or treated stereotypically, or with the prospect of conforming to the stereotype.

Diese Bedrohung wird als situationsbezogen charakterisiert und kann bei Mitgliedern stereotypisierter Gruppen (z.B.: Frauen), die sich mit der Domäne (z.B.: Mathematik), in der sich ein negatives Stereotyp anwenden lässt, identifizieren, zu einer Leistungsminderung führen (Steele, 1997).

Gemäß den AutorInnen kann Stereotype Threat zudem bewirken, dass sich stereotypisierte Personen nicht mehr mit der jeweiligen Domäne (z.B.: Mathematik) identifizieren und andere Fachgebiete (z.B.: Englisch) heranziehen, auf die sie ihre Identität stützen können (Spencer et al., 1999; Steele, 1997).

3.3 Auswirkungen auf die Leistung

Im Rahmen der ersten empirischen Studien zum Phänomen Stereotype Threat wurden die Auswirkungen negativer kultureller Stereotypen auf die Testleistung analysiert (Campbell & Collaer, 2009).

In einer Untersuchung von Steele und Aronson (1995) wurde afroamerikanischen und weißen StudentInnen ein schwieriger verbaler Test vorgegeben. In einer Bedingung wurde den StudentInnen vor der Testbearbeitung der Test als ein Verfahren dargestellt, der ihre verbalen Fähigkeiten abbilde (*Stereotype Threat-Bedingung*). Den StudentInnen in der zweiten Bedingung hingegen, wurde derselbe Test als Problemlöseaufgabe beschrieben und kein Bezug zu den verbalen Fähigkeiten hergestellt (*Non-Threat-Bedingung*). Die Autoren vermuteten, dass die Darstellung des

Tests in der *Stereotype Threat-Bedingung* bei den afroamerikanischen StudentInnen das negative kulturelle Stereotyp (geringe intellektuelle Leistungsfähigkeit) relevant macht und sich die Bedrohung im Hinblick auf eine etwaige Bestätigung des Stereotyps, negativ auf die Testleistung auswirkt. In Einklang mit den Erwartungen zeigte sich, dass in der *Stereotype Threat-Bedingung* die afroamerikanischen StudentInnen signifikant schlechtere Leistungen als die weißen StudentInnen erzielten, hingegen in der *Non-Threat-Bedingung* erbrachten die afroamerikanischen StudentInnen eine ebenbürtige Leistung wie die weißen StudentInnen.

In Bezug auf die Leistungsfähigkeit von Frauen, konnten die negativen Auswirkungen von Stereotype Threat beispielsweise bei schwierigen Mathematikaufgaben (z.B.: Spencer et al., 1999) sowie visuell-räumlichen Aufgaben (Campbell & Collaer, 2009) verdeutlicht werden, wie auch später dargelegt wird.

Zudem konnte die Leistungsminderung sogar bei Testpersonen, die hingegen nicht von negativen Stereotypen im Hinblick auf die untersuchte Domäne betroffen sind, von Aronson et al. (1999) aufgezeigt werden. In der Untersuchung wurden weiße, männliche Studenten mit besonders hohen mathematischen Fähigkeiten in einer Bedingung sowohl verbal als auch anhand von Artikeln vermittelt, dass asiatische Studenten in mathematischen Domänen die weißen Studenten übertreffen würden (*Stereotype Threat-Bedingung*). Die Testergebnisse bei einem schwierigen Mathematiktest zeigten, dass die Studenten in der *Stereotype Threat-Bedingung* eine geringere Leistung im Vergleich zu Studenten, die nicht auf ein diesbezügliches Stereotyp hingewiesen wurden, erreichten.

Im Folgenden wird auf Stereotype Threat im Zusammenhang mit Leistungsfähigkeit bei Frauen näher eingegangen und exemplarische Untersuchungen dargestellt.

Spencer et al. (1999) beleuchteten die Auswirkungen von Stereotype Threat auf die Leistungen von StudentInnen in einem schwierigen Mathematiktest. In der Studie wurde die Relevanz des Geschlechtsstereotyps „Männer sind in Mathematik besser als Frauen“ variiert, indem die UntersuchungsteilnehmerInnen in einer Bedingung explizit darauf hingewiesen wurden, dass es bei dem Mathematiktest in der Vergangenheit Geschlechtsunterschiede gegeben habe (*Stereotype Threat-Bedingung*), hingegen in der *Non-Threat-Bedingung* wurde bekundet, dass man bei dem Mathematiktest in der Vergangenheit nie Geschlechtsunterschiede festgestellt habe. Die AutorInnen kamen zu dem Ergebnis, dass Frauen in der *Stereotype Threat-Bedingung* geringere Leistungen als Männer erzielten, während in der *Non-Threat-Bedingung* Frauen

genauso gute Leistungen wie Männer erbrachten. Die Untersuchungsergebnisse weisen darauf hin, dass Mathematikleistungen von Frauen durch die Entfernung bzw. Reduzierung der Bedrohung durch Stereotype, begünstigt werden können.

In einer Untersuchung von Smith und White (2002) wurde 70 Studentinnen ein Mathematiktest computerbasiert vorgegeben und drei experimentelle Manipulationen verwendet. In einer Bedingung wurde den Studentinnen vermittelt, dass Männer in Mathematik besser seien als Frauen (*explizite Aktivierung* von Stereotype Threat), wohingegen in der zweiten Bedingung nichts Diesbezügliches gesagt wurde (*implizite Aktivierung* von Stereotype Threat). In der dritten Bedingung (*nullified Stereotyp*) erfolgte die Entkräftung des negativen Stereotyps, indem bekundet wurde, dass Frauen und Männer gleich gute Leistungen in Mathematik erbrachten. Die Ergebnisse verdeutlichten, dass die Studentinnen in den Bedingungen *explizite Aktivierung* von Stereotype Threat und *implizite Aktivierung* von Stereotype Threat geringere Leistungen erzielten als die Studentinnen in der *nullified* Bedingung. Demgemäß scheint die Platzierung von Frauen in eine relevante Testsituation hinreichend zu sein, um Stereotype Threat-Effekte hervorzurufen.

Dass sich dieser Befund nicht nur auf die Mathematikleistung von Frauen, sondern auch auf die visuell-räumliche Leistungsfähigkeit bezieht, konnten Campbell und Collaer (2009) demonstrieren. Die Autorinnen setzten die drei experimentellen Manipulationen von Stereotype Threat basierend auf Smith und White (2002), adaptiert im Hinblick auf die visuell-räumliche Leistungsfähigkeit von StudentInnen, ein (z.B.: *explizite Aktivierung* „Männer erzielen bei visuell-räumlichen Aufgaben eine höhere Leistung als Frauen“). Im Unterschied zur vorherigen Studie wurden nicht nur weibliche, sondern auch männliche Studenten in den Untersuchungsplan einbezogen. Die Ergebnisse bei einem visuell-räumlichen Fähigkeitstest zeigten, dass Frauen in der *nullified* Bedingung eine den Männern ebenbürtige Leistung erzielten, während sie in den beiden anderen Bedingungen (*explizit* und *implizite Aktivierung*) eine geringere Leistung als Männer erreichten. Dieser Befund deutet darauf hin, dass bei Frauen das negative Stereotyp im Hinblick auf die männliche Überlegenheit in dieser Domäne, alleinig durch die Präsentation der visuell-räumlichen Aufgaben aktiviert wurde.

Laut Keller (2007) sowie Keller und Dauenheimer (2003) ist fraglich, ob Stereotype Threat-Effekte auch in einem natürlichen Setting feststellbar sind, da die überwiegende Mehrheit der Untersuchungen zu diesem Phänomen im Laborsetting durchgeführt wurde und man die UntersuchungsteilnehmerInnen meist nach bestimmten Kriterien

(z.B.: besonders hohe mathematische Fähigkeiten, hohe Identifikation mit der Leistungsdomäne) auswählte.

In diesem Zusammenhang wurde Stereotype Threat in einem natürlichen Setting (im Klassenzimmer) sowie ohne eine Auswahl von SchülerInnen nach bestimmten Kriterien, von Keller und Dauenheimer (2003) untersucht: SchülerInnen mit einem Durchschnittsalter von 15.7 Jahren wurden in einer Bedingung darauf hingewiesen, dass es bei dem Mathematiktest in der Vergangenheit Geschlechtsunterschiede gegeben habe (*Stereotype Threat-Bedingung*) und in der *Non-Threat-Bedingung* wurde das Gegenteil bekundet. Die Autoren fanden, dass Mädchen in der *Non-Threat-Bedingung* eine höhere Leistung als Mädchen in der *Stereotype Threat-Bedingung* erzielten. Der Vergleich der Geschlechtergruppen ergab zunächst in der *Stereotype Threat-Bedingung*, dass Mädchen eine geringere Leistung als Burschen aufwiesen, jedoch in der *Non-Threat-Bedingung* die Mädchen ebenso gute Leistungen wie Burschen erbrachten.

Des Weiteren zeigte eine Studie von Keller (2007) im natürlichen Setting, dass in der *Non-Threat-Bedingung* Mädchen mit einer hohen Domänenidentifikation (z.B.: persönliche Wichtigkeit in Mathematik gut zu sein) eine höhere Mathematikleistung erzielten als Mädchen mit einer niedrigen Domänenidentifikation, während in der *Stereotype Threat-Bedingung* das Gegenteil der Fall war. Im Hinblick auf die Geschlechtergruppen konnten allerdings in der *Stereotype Threat-Bedingung* keine Unterschiede in der Mathematikleistung zwischen den Mädchen und Burschen aufgezeigt werden. Der Autor führt dies auf das Alter ($M = 15.9$ Jahre) der getesteten SchülerInnen zurück, wonach sich größere Leistungsunterschiede zwischen Mädchen und Burschen in Bezug auf die Mathematik vorwiegend erst zu einem späteren Zeitpunkt abbilden.

Dies entspricht auch den Untersuchungsergebnissen einer Metaanalyse von Hyde, Fennema & Lamon (1990, zitiert nach Steele, 1997, S. 615), auf die bereits in der *Einführung* eingegangen wurde (siehe S. 8).

Insgesamt weisen die Ergebnisse darauf hin, dass Stereotype Threat auch in einem natürlichen Setting, wie im Klassenzimmer, vorherrscht (Keller, 2007; Keller & Dauenheimer, 2003).

3.4 Vermittelnde Mechanismen

Die negativen Auswirkungen von Stereotype Threat auf die Leistung (siehe vorigen Abschnitt) konnten in vielen Studien verdeutlicht werden (z.B.: Spencer et al., 1999). In diesem Sinne ist die Darlegung effektiver Interventionen von der Eruiierung der vermittelnden Mechanismen der Leistungsverschlechterung unter Stereotype Threat abhängig (Keller & Dauenheimer, 2003). Allerdings ist im Rahmen der Vielzahl an durchgeführten Untersuchungen nicht belegt, welche Mechanismen genau beteiligt sind (Cadinu, Maass, Rosabianca & Kiesner, 2005; Keller & Dauenheimer, 2003).

In diesem Abschnitt werden ausgewählte Mechanismen, die die Leistungsminderung von Frauen unter Stereotype Threat bewirken könnten, dargestellt.

Laut Steele (1997) resultiert die Leistungsminderung unter Stereotype Threat aufgrund einer emotionalen Reaktion, die aus der wahrgenommenen Bedrohung hervorgeht.

In diesem Zusammenhang untersuchten Spencer et al. (1999) unter anderem die Ängstlichkeit als etwaige Mediatorvariable der Leistungsverringerng von Frauen unter Stereotype Threat. In der Studie konnten medierende Effekte von Angst auf die Leistungsminderung unter Stereotype Threat beobachtet werden. An dieser Stelle ist darauf hinzuweisen, dass die Ängstlichkeit vor der Bearbeitung der mathematischen Aufgaben erhoben wurde.

Bei Keller und Dauenheimer (2003), die die Ängstlichkeit nach der Vorgabe der mathematischen Aufgaben erfassten, konnte kein Einfluss der Stereotype Threat Manipulation auf die Ängstlichkeit festgestellt werden.

Neue Erkenntnisse kommen von Marx und Stapel (2006): 133 StudentInnen wurde ein schwieriger Mathematiktests vorgegeben und die Ängstlichkeit sowohl vor als auch nach der Testbearbeitung erhoben. Die Ergebnisse zeigten, dass Frauen in der *Stereotype Threat-Bedingung* eine höhere Ängstlichkeit vor der Testbearbeitung angaben als nachher. In der Studie konnte somit verdeutlicht werden, dass Stereotype Threat eine erhöhte Ängstlichkeit bei Frauen bewirken kann, jedoch vor der Testung relevant ist. In Bezug auf die Mathematikleistung konnte jedoch kein starker Zusammenhang zwischen dem Ängstlichkeitslevel und der Testleistung beobachtet werden.

Darüber hinaus erfolgte in dieser Studie auch die Erfassung der Frustration: Frauen in der *Stereotype Threat-Bedingung* berichteten über ein höheres Frustrationslevel nach

der Bearbeitung des Mathematiktests als vorher, zudem gaben sie nach der Bearbeitung auch ein höheres Frustrationslevel an im Vergleich zu Männern. Im Hinblick auf die Mathematikleistung stellten die Autoren bei den Frauen in der *Stereotype Threat-Bedingung* fest, dass das höhere Frustrationslevel nach der Testbearbeitung mit einer geringeren Mathematikleistung verbunden ist. Gemäß den Untersuchungsergebnissen scheint die erhöhte Ängstlichkeit bei Frauen unter Stereotype Threat auf die Zeit vor der Testung beschränkt zu sein und wird von Frustration nach der Mathematiktestbearbeitung ersetzt (Marx & Stapel, 2006).

Cadinu et al. (2005) untersuchten als etwaigen Mediator der Leistungsminderung von Frauen unter Stereotype Threat: Störende Gedanken. In der Studie wurden Psychologiestudentinnen schwierige Mathematikaufgaben vorgegeben und ihnen mitgeteilt, dass sie auf das jeweils leere Blatt vor jeder Aufgabe alles niederschreiben sollen, was ihnen gerade in den Sinn kommt. Gemäß den Ergebnissen berichteten die Studentinnen in der *Stereotype Threat-Bedingung* häufiger negative Gedanken hinsichtlich mathematischer Inkompetenz und Leistungsunfähigkeit als ihre Kolleginnen in der *Non-Threat-Bedingung*. Darüber hinaus konnte bei den Studentinnen in der *Stereotype Threat-Bedingung* eine deutlich geringere Mathematikleistung festgestellt werden, die durch die negativen Gedanken vermittelt wurde. Laut AutorInnen bedingen die negativen Gedanken die Leistungsminderung (von Frauen unter Stereotype Threat) und nicht umgekehrt, wobei aber dieser Faktor nicht den einzigen Mechanismus darstellen muss, auf dem diese Leistungsverringerung basieren könnte.

In einer Untersuchung von Cadinu, Maass, Frigerio, Impagliazzo und Latinotti (2003) wurde die Leistungserwartung beleuchtet. Die Autorinnen gingen davon aus, dass Frauen unter Stereotype Threat eine geringere Mathematikleistung erbringen, da sie womöglich niedrige Erwartungen in Bezug auf ihre eigenen Mathematikleistungen aufweisen. An der Studie nahmen insgesamt 95 Psychologiestudentinnen teil, die vor der Bearbeitung schwieriger Mathematikaufgaben auf einer Grafik ihre Leistungserwartungen in Bezug auf ihre mathematisch-logischen Fähigkeiten einschätzten. Die Autorinnen stellten fest, dass Studentinnen in der *Stereotype Threat-Bedingung* eine niedrigere Leistungserwartung aufwiesen sowie eine geringere Mathematikleistung erzielten, wobei sich dieses Ergebnis aber nur auf diejenigen Studentinnen bezog, die sich mit Mathematik stark identifizieren. Demnach schätzten die Studentinnen in dieser Bedingung ihre mathematisch-logischen Fähigkeiten gering ein und erzielten tatsächlich niedrige Leistungen. In der Studie konnte die

Leistungserwartung als partieller Mediator der Stereotype Threat-Effekte ermittelt werden.

Hingegen bei Sekaquaptewa und Thompson (2003) konnte dieser Befund nicht bestätigt werden. In der Studie schätzten PsychologiestudentInnen ihre Leistungserwartung auf einer 10-Punkte-Skala ein, bevor sie Fragen, basierend auf davor bereitgestellten mathematischen Informationen, beantworten sollten. Gemäß den Ergebnissen, kann die Leistungserwartung nicht als Mediator der Leistungsminderung von Frauen unter Stereotype Threat angenommen werden. Die Autorinnen vermuten, dass die Leistungsminderung auch aus impliziten Mechanismen hervorgehen könnte.

3.5 Leistungsminderungen reduzieren

Bezugnehmend auf die dargelegten Studien (vgl. Unterpunkt „*Auswirkungen auf die Leistung*“) konnte gezeigt werden, dass in der *Non-Threat-Bedingung* (z.B.: Spencer et al., 1999; Steele & Aronson, 1995) respektive *nullified Stereotyp* (Bedingung) (z.B.: Smith & White, 2002) keine Leistungsminderung von Frauen bzw. AfroamerikanerInnen beobachtet wurde. Demnach zeigte sich eine Milderung der negativen Auswirkungen von Stereotype Threat, indem die Leistung nicht in Beziehung zu den getesteten Fähigkeiten gebracht wurde (Steele & Aronson, 1995) oder die Testleistung als geschlechtsunabhängig dargestellt wurde (z.B.: Smith & White, 2002). Darüber hinaus erfolgte die Beleuchtung vieler Strategien, von denen man vermutete, dass sie die Leistungsverschlechterung unter Stereotype Threat reduzieren könnten (z.B.: McIntyre, Paulson & Lord, 2003). Im weiteren Abschnitt erfolgt die Darlegung von vier bestimmten Strategien.

Johns, Schmader und Martens (2005) überprüften als mögliche Intervention die Aufklärung über die Effekte von Stereotype Threat: 144 StudentInnen wurden drei Bedingungen (*Stereotype Threat-Bedingung*, *Non-Threat-Bedingung* und *Interventions-Bedingung*) randomisiert zugewiesen und anspruchsvolle mathematische Aufgaben vorgegeben. In der *Interventions-Bedingung* wurde die gleiche Instruktion wie in der Stereotype Threat-Bedingung verwendet, aber darüber hinaus wurden die StudentInnen über das Phänomen Stereotype Threat informiert und ihnen vermittelt, dass Ängstlichkeit im Zuge der Testbearbeitung mit negativen Stereotypen zusammenhänge und nichts über die gegenwärtigen Fähigkeiten aussage. Laut AutorInnen könnte die Beschreibung den Frauen eine Erklärungsmöglichkeit für

potentiell auftretende Ängstlichkeit während der Testbearbeitung darstellen. Die Studienergebnisse zeigten zunächst in der *Interventions-Bedingung*, dass Frauen eine ebenbürtige Mathematikleistung wie Männer erzielten. Zudem erreichten Frauen in der *Interventions-Bedingung* eine ebenso gute Leistung wie Frauen in der *Non-Threat-Bedingung*. In der *Stereotype Threat-Bedingung* hingegen, erbrachten Frauen eine deutlich geringere Leistung als Frauen in der *Interventions-Bedingung*. Demnach scheint die Aufklärung über Stereotype Threat-Effekte eine praktikable Technik darzustellen, um die negativen Auswirkungen zu mildern.

Bei McIntyre et al. (2003) folgte kurz nach der Stereotype Threat Induzierung („Männer sind in Mathematik besser als Frauen“) das Lesen sowie die kritische Auseinandersetzung mit biographischen Porträts über erfolgreiche Frauen in männerdominierenden Berufsfeldern (Medizin, Architektur, Rechtswissenschaften und Invention). In der anderen Bedingung bezogen sich die Porträts auf erfolgreiche Unternehmen. Gemäß den Ergebnissen erreichten weibliche Untersuchungsteilnehmer, die sich mit Porträts erfolgreicher Frauen auseinandersetzten, eine bessere Mathematikleistung als weibliche Testpersonen in der „Unternehmens“-Bedingung sowie eine ebenso gute Leistung wie männliche Testpersonen in derselben Bedingung. Demnach kann die Leistungsminderung unter Stereotype Threat gelindert werden, wenn Leistungen bzw. Erfolge anderer Frauen in Erinnerung gebracht werden.

Einen leistungsfördernden Effekt zeigte auch die Darlegung von kompetenten weiblichen Rollenvorbildern, wie Marx und Roman (2002) feststellten. In der Studie erzielten Frauen in einem schwierigen Mathematiktest eine ebenso gute Leistung wie Männer, wenn die potentielle Versuchsleiterin (diese war nicht präsent) als besonders kompetent in Mathematik dargestellt wurde. Die gegenteilige Beschreibung der potentiellen Versuchsleiterin führte bei Frauen zu einer geringeren Mathematikleistung im Vergleich zu Männern. Die Ergebnisse weisen darauf hin, dass Beispiele von Frauen, die dem Geschlechtsstereotyp widersprechen, die Mathematikleistung von Frauen begünstigen können.

In einer Untersuchung von Good, Aronson und Inzlicht (2003) wurden 138 SchülerInnen im Rahmen eines Trainings verschiedene Perspektiven hinsichtlich der schulischen Leistungsfähigkeit vermittelt. Die AutorInnen untersuchten, ob das jeweilige Training Stereotype Threat-Effekte reduzieren kann. Schülerinnen, die von StudentInnen (MentorInnen) ein Schuljahr dahingehend unterrichtet und bekräftigt

wurden, dass Intelligenz durch mentale Arbeit gestärkt werden kann und/oder Leistungsschwierigkeiten hinsichtlich des Übergangs in eine höhere Schule auf die neue Situation rückführbar seien, erzielten eine höhere Mathematikleistung als Schülerinnen in der Kontrollgruppe (hier ging es um die negativen Auswirkungen von Drogen). Demgemäß scheint die Vermittlung von differenzierten Einstellungen, in Bezug auf die schulische Leistungsfähigkeit, den leistungsmindernden Effekt von Stereotype Threat zu hemmen.

4. Selbstwirksamkeitserwartung

4.1 Begriffsbestimmung

Das Konzept der Selbstwirksamkeitserwartung basiert auf Albert Bandura, der es 1977 im Rahmen eines wissenschaftlichen Artikels erstmals skizzierte (Bandura, 1977). In der sozial kognitiven Theorie wird der Selbstwirksamkeitserwartung, im Zusammenspiel mit anderen Mechanismen, eine zentrale Bedeutung in der Determinierung des menschlichen Verhaltens zugesprochen (Bandura, 1986, 1997).

Gemäß Bandura (1986) versteht man unter dem Begriff *Selbstwirksamkeitserwartung* „people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances“ (S. 391). Selbstwirksamkeitserwartungen basieren somit nicht auf den eigentlichen Fähigkeiten einer Person, sondern auf der persönlichen Beurteilung (Bandura, 1986).

Schwarzer und Jerusalem (2002) definieren *Selbstwirksamkeitserwartungen* „als die subjektive Gewissheit, neue oder schwierige Anforderungssituationen auf Grund eigener Kompetenz bewältigen zu können“ (S. 35).

In diesem Sinne sind die Anforderungssituationen dadurch charakterisiert, dass deren Bewältigung sowohl Anstrengung als auch Ausdauer bedürfen. Routinemäßige Aufgaben werden hierbei nicht miteinbezogen (Jerusalem & Mittag, 1999; Schwarzer & Jerusalem, 2002).

Der Begriff Selbstwirksamkeitserwartung wird von *Konsequenzerwartungen* (*outcome expectations*) abgegrenzt (siehe Abbildung 1): Während sich erstes auf die persönliche Überzeugung bezieht, ein bestimmtes Verhalten, das für ein bestimmtes Ergebnis erforderlich ist, ausführen zu können, steht bei zweitem die Beurteilung, der mit einem bestimmten Verhalten, verbundenen *Konsequenzen* im Vordergrund. Ein Beispiel für Selbstwirksamkeitserwartungen stellt die persönliche Überzeugung einer Person hinsichtlich der eigenen Fähigkeiten dar, vor einem großen Publikum eine Rede halten zu können. Die mit dem jeweiligen Verhalten erwarteten Konsequenzen (z.B.: Applaus) kennzeichnen Konsequenzerwartungen. In diesem Zusammenhang können niedrige Selbstwirksamkeitserwartungen (Zweifel, ob man das erforderliche Verhalten

überhaupt ausführen kann) eine Person hindern im Sinne der Konsequenzerwartungen zu handeln (Bandura, 1977, 1986, 1997).

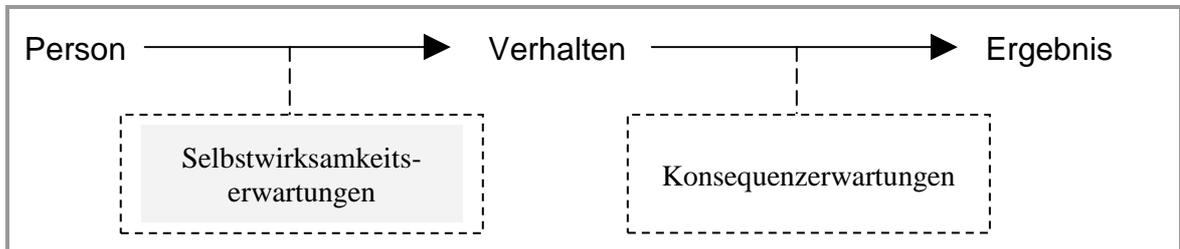


Abbildung 1: Abgrenzung zwischen Selbstwirksamkeitserwartungen und Konsequenzerwartungen (Bandura, 1977, S. 193)

Selbstwirksamkeitserwartungen stellen persönliche Ressourcen dar (Schwarzer, 2004). Sie sind entscheidend „für hohe Motivation und hohes Leistungsniveau, für psychisches und körperliches Wohlbefinden und für hohe Berufs- und Lebenszufriedenheit“ (Schwarzer, 2004, S. 13).

Personen mit starken bzw. *hohen* Selbstwirksamkeitserwartungen sind motiviert sich neuen, herausfordernden Aufgaben zu stellen; sie zeigen ein hohes Maß an Anstrengungsbereitschaft sowie Ausdauer und somit steigt auch die Wahrscheinlichkeit auf Erfolg (Bandura, 1977, 1986, 1997; Jerusalem, 2002; Jerusalem & Mittag, 1999). Im Gegensatz dazu wagen sich Personen mit *niedrigen* Selbstwirksamkeitserwartungen oftmals nicht an die jeweiligen Aufgaben heran respektive zeigen sie ein vorzeitiges Aufgeben (Bandura, 1977, 1986, 1997; Jerusalem, 2002). Bezugnehmend auf die berufliche Entwicklung, kann ein positiver Zusammenhang zwischen der Selbstwirksamkeitserwartung und den beruflichen Optionen, die eine Person in Betracht zieht, angenommen werden (Bandura, 1986, 1997). Personen mit *niedrigen* Selbstwirksamkeitserwartungen weisen zudem Unentschlossenheit hinsichtlich der (geringen) erwägten Karrieremöglichkeiten auf (Taylor & Betz, 1983, zitiert nach Bandura, 1986, S. 432).

4.2 Struktur der Selbstwirksamkeitserwartungen

Gemäß Bandura (1977, 1986, 1997) können Selbstwirksamkeitserwartungen anhand von drei Dimensionen unterschieden werden: Die erste Dimension stellt das *Ausmaß* (magnitude, level) der Selbstwirksamkeitserwartungen dar und kennzeichnet, ob sich Selbstwirksamkeitserwartungen auf leichte, mittelschwere oder auch schwierigere

Anforderungen beziehen. Demnach kann eine Person zum Beispiel davon überzeugt sein einen Zehn-Kilometer-Lauf zu meistern, aber sich die Kompetenz hinsichtlich der erfolgreichen Bewältigung eines Halbmarathons nicht zuschreiben.

Der *Generalitätsgrad* (generality) der Selbstwirksamkeitserwartungen bildet die zweite Dimension: Eine Person kann sich beispielsweise ein hohes Maß an Selbstwirksamkeitserwartung hinsichtlich einer bestimmten Anforderung oder in einem bestimmten Bereich (z.B.: Mathematik) zuschreiben, jedoch in einer anderen Domäne (z.B.: Deutsch) oder im Rahmen des allgemeinen Aktivitätsspektrums als gering selbstwirksam einschätzen (Bandura, 1977, 1986, 1997). Die *allgemeine* Selbstwirksamkeitserwartung richtet sich auf die Beurteilung der allgemeinen Fähigkeit zur Lebensbewältigung (Schwarzer, 2004; Schwarzer & Jerusalem, 2002). Ein Itembeispiel der *Skala zur Allgemeinen Selbstwirksamkeitserwartung*² von Jerusalem und Schwarzer (1999) lautet: „Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen“. Auf ein bestimmtes Feld (z.B.: Ernährung, Schule) bezieht sich die *bereichsspezifische* Selbstwirksamkeitserwartung (Schwarzer, 2004; Schwarzer & Jerusalem, 2002). In diesem Zusammenhang stellen unter anderem die Skala zur Messung der *schulbezogenen Selbstwirksamkeitserwartung* von Jerusalem und Satow (1999) (Itembeispiel der Skala: „Selbst wenn ich mal längere Zeit krank sein sollte, kann ich immer noch gute Leistungen erzielen“) oder die Skala zur Messung der *Selbstwirksamkeitserwartung im Umgang mit sozialen Anforderungen* von Satow und Mittag (1999) (Itembeispiel der Skala: „Ich traue mich zu sagen, was ich denke, auch wenn die anderen nicht meiner Meinung sind“), Erhebungsinstrumente zur Erfassung bereichsspezifischer Selbstwirksamkeitserwartung dar. *Situationsspezifische* Selbstwirksamkeitserwartungen richten sich auf die Bewältigung einer bestimmten Handlung angesichts bestimmter Hindernisse (z.B.: Ich bin mir sicher, dass ich am Nachmittag für die Prüfung lerne, auch wenn Freunde mich zum Snowboarden einladen) (Schwarzer, 2004; Schwarzer & Jerusalem, 2002).

Laut Schwarzer, Bäßler, Kwiatek, Schröder und Zhang (1997) wurden Selbstwirksamkeitserwartungen vorwiegend domänenspezifisch interpretiert. In diesem Zusammenhang beziehen sich viele Erhebungsinstrumente hinsichtlich Selbstwirksamkeitserwartungen auf bestimmte Bereiche bzw. bestimmte Situationen.

² In der vorliegenden Diplomarbeit wurde die *allgemeine* Selbstwirksamkeitserwartung erhoben. Die Darstellung des eingesetzten Erhebungsinstruments *Skala zur Allgemeinen Selbstwirksamkeitserwartung* von Jerusalem und Schwarzer (1999) ist dem methodischen Teil (Unterpunkt 7.2.6) zu entnehmen.

Neben der *Skala zur Allgemeinen Selbstwirksamkeitserwartung* von Jerusalem und Schwarzer (1999) gibt es jedoch auch weitere Autoren (z.B.: Sherer & Maddux, 1982, zitiert nach Schwarzer et al., 1997, S. 71), die generalisierte Formen konstruierten.

Die dritte Dimension stellt schließlich die *Stärke* (strength) der Selbstwirksamkeitserwartungen dar: Während Personen mit niedrig ausgeprägten Selbstwirksamkeitserwartungen angesichts von Hindernissen oder Rückschlägen dazu tendieren aufzugeben, begünstigen starke bzw. beständige Selbstwirksamkeitserwartungen in diesem Kontext, dass beharrlich durchgehalten wird (Bandura, 1977, 1986, 1997).

4.3 Quellen der Selbstwirksamkeitserwartungen

Der Aufbau von Selbstwirksamkeitserwartungen resultiert nach Bandura (1977, 1986, 1997) hauptsächlich aus vier bestimmten Quellen. In diesem Unterpunkt werden die Quellen hinsichtlich ihrer Einflussstärke absteigend dargestellt.

Die erste und somit einflussreichste Quelle bildet die *persönliche (Erfolgs-) Erfahrung* (enactive mastery experience bzw. enactive attainment). Im Allgemeinen führen persönliche Erfolge zu einer Steigerung der Selbstwirksamkeitserwartungen und Misserfolge bewirken den gegenteiligen Effekt. Allerdings ist in diesem Zusammenhang die Ausprägung der bereits bestehenden Selbstwirksamkeitserwartungen hinsichtlich der persönlichen Gewichtung der jeweiligen Erfolgs- oder Misserfolgserfahrung von Relevanz: Demnach kann beim Auftreten vereinzelter Misserfolge bei Personen mit stark ausgeprägten Selbstwirksamkeitserwartungen keine beträchtliche Verringerung der Selbstwirksamkeitserwartungen angenommen werden im Vergleich zu Personen mit niedrigen Selbstwirksamkeitserwartungen. Für die Stärkung respektive Schwächung der Selbstwirksamkeitserwartungen sind zum Beispiel die investierte Anstrengung und der Schwierigkeitsgrad der jeweiligen Anforderung ebenso von Bedeutsamkeit. Auch das Überwinden von Misserfolgen kann zur Stärkung von Selbstwirksamkeitserwartungen beitragen (Bandura, 1977, 1986, 1997).

Die zweite Quelle zum Aufbau von Selbstwirksamkeitserwartungen stellt laut Bandura (1977, 1986, 1997) die *stellvertretende Erfahrung* (vicarious experience) dar. Im Sinne von: „Wenn andere das schaffen, schaffe ich das auch“, können Selbstwirksamkeitserwartungen gestärkt werden, indem man eine ähnliche Person bei

der erfolgreichen Meisterung einer Tätigkeit beobachtet. Die gegenteilige Beobachtung führt meist zu einer Schwächung der Selbstwirksamkeitserwartungen. Die Stärkung respektive Schwächung von Selbstwirksamkeitserwartungen anhand von stellvertretenden Erfahrungen resultiert hauptsächlich aus sozialen Vergleichsprozessen. Ein weiterer bedeutender Faktor bildet in diesem Zusammenhang das mögliche Aneignen effektiver Strategien zur Meisterung von Herausforderungen aufgrund der Demonstration der Modellperson.

Die *verbale Überzeugung* (verbal persuasion) kennzeichnet die dritte Quelle, hierbei steht die Meinung einer anderen Person im Fokus. Demnach können Selbstwirksamkeitserwartungen gestärkt werden, indem man von einer anderen Person dahingehend verbal überzeugt wird, dass man für die erfolgreiche Bewältigung einer bestimmten Anforderung, die dafür erforderlichen Kompetenzen verfügt. Dabei ist von Relevanz, dass die mittels verbaler Überzeugung gestärkten Selbstwirksamkeitserwartungen in einem realistischen Rahmen liegen, da ansonsten Misserfolge begünstigt werden und somit eine Schwächung der Selbstwirksamkeitserwartungen einhergehen kann. Laut Autor kann eine dauerhafte Stärkung der Selbstwirksamkeitserwartungen nur aufgrund verbaler Überzeugung nicht herbeigeführt werden, da widersprechende Erfahrungen die Selbstwirksamkeitserwartungen schwächen würden. Die verbale Überzeugung beeinflusst jedoch auch die Anstrengungsbereitschaft des Empfängers: Während die „positive“ verbale Überzeugung die Anstrengungsbereitschaft einer Person stärkt, führt die „negative“ verbale Überzeugung (z.B.: „Du schaffst das nicht“) zur Vermeidung von Herausforderungen und begünstigt vorzeitiges aufgeben. An dieser Stelle ist darauf hinzuweisen, dass die Stärkung respektive Schwächung der Selbstwirksamkeitserwartungen anhand verbaler Überzeugung auch von der subjektiven Einschätzung des Empfängers hinsichtlich der Glaubwürdigkeit sowie Fachkundigkeit des „Überzeugers“ abhängt (Bandura, 1977, 1986, 1997).

Die vierte Quelle stellt gemäß Bandura (1977, 1986, 1997) *die emotionale Erregung* (emotional arousal bzw. physiological state) dar. Diese kann von einer Person zur Beurteilung ihrer persönlichen Kompetenz bzw. Inkompetenz herangezogen werden. Im Allgemeinen wird starke emotionale Erregung mit Misserfolgserwartung in Verbindung gebracht und niedrige emotionale Erregung mit Erwartung auf Erfolg. Für das Ausmaß der Stärkung respektive Schwächung der Selbstwirksamkeitserwartungen sind hierbei jedoch mehrere Aspekte in Bezug auf die emotionalen Erregung (z.B.: die jeweiligen Umstände) sowie deren persönliche Interpretation relevant.

Aufgrund vergangener Erfahrungen kann beispielsweise emotionale Erregung von einer Person als leistungsfördernd oder leistungshemmend interpretiert werden und somit Selbstwirksamkeitserwartungen stärken bzw. schwächen (Bandura, 1986, 1997).

4.4 Selbstwirksamkeitserwartung und Leistung

Eine Vielzahl an Studien beschäftigt sich mit der Untersuchung des Zusammenhangs zwischen dem Konstrukt der Selbstwirksamkeitserwartung und Leistung (z.B.: Pajares & Miller, 1994). Nachfolgend werden ausgewählte Untersuchungen dieses Forschungsfeldes dargestellt und auf geschlechtsspezifische Unterschiede in der Selbstwirksamkeitserwartung eingegangen.

Luszczynska, Gutiérrez-Doña und Schwarzer (2005) analysierten anhand einer Stichprobe von insgesamt 8796 UntersuchungsteilnehmerInnen (42,7 % Männer und 57,3 % Frauen) aus fünf verschiedenen Ländern (Deutschland, Costa Rica, Polen, Türkei und Amerika) den Zusammenhang zwischen der Allgemeinen Selbstwirksamkeitserwartung und interessierenden Variablen. Das Lebensalter der Stichprobe lag zwischen 13 und 77 Jahren, wobei die Stichprobe mehrheitlich SchülerInnen bzw. StudentInnen umfasste. Bezugnehmend auf die Leistung zeigten die Ergebnisse, dass Personen mit einer hohen Selbstwirksamkeitserwartung höhere schulische Leistungen erzielen als jene mit einer niedrigen Selbstwirksamkeitserwartung. Zudem stellten die AutorInnen einen positiven Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Zufriedenheit in der Schule bzw. im Beruf fest.

Eine Metaanalyse von Multon, Brown und Lent (1991) belegt ebenso den positiven Zusammenhang zwischen der Selbstwirksamkeitserwartung und Leistung: 36 Studien mit insgesamt 4998 Testpersonen wurden in die Untersuchung einbezogen, wobei das Durchschnittsalter der UntersuchungsteilnehmerInnen 16.6 Jahre betrug. Die in den Studien eingesetzten Erhebungsinstrumente wurden einer von drei Kategorien zugeordnet: Standardisierter Leistungstest, schulbezogene Erhebungsmethoden und fähigkeitsspezifische Aufgaben. Die Mehrheit der Studien kennzeichnet die dritte Kategorie. Die Berechnung der Effektgröße des Zusammenhangs zwischen der Selbstwirksamkeitserwartung und der Leistung ergab $r = 0.38$ und weist unter Berücksichtigung der Effektgrößenklassifikation (vgl. Bortz & Döring, 2005, S. 604) auf einen mittleren Effekt hin.

Pajares und Miller (1994) gaben 350 StudentInnen (229 Frauen und 121 Männer) mathematische Problemlöseaufgaben vor und erhoben neben der mathematikbezogenen Selbstwirksamkeitserwartung fünf weitere Variablen (wahrgenommene Nützlichkeit von Mathematik, Ängstlichkeit in Bezug auf Mathematik, mathematisches Selbstkonzept und vorausgegangene mathematische Erfahrung zum einen in Bezug auf die Sekundarstufe und zum anderen hinsichtlich der Hochschule). Anhand einer Pfadanalyse konnte die Selbstwirksamkeitserwartung als stärkster Prädiktor der Mathematikleistung bei den StudentInnen identifiziert werden. Geschlechtsspezifische Unterschiede zugunsten der männlichen Untersuchungsteilnehmer zeigten sich sowohl bei der Selbstwirksamkeitserwartung als auch bei der Mathematikleistung. Im Rahmen der Untersuchung konnte die Selbstwirksamkeitserwartung als Mediator des Effekts von Geschlecht auf die Mathematikleistung eruiert werden.

In einer ähnlichen Studie von Pajares und Kranzler (1995) konnte die mathematikbezogene Selbstwirksamkeitserwartung neben der allgemeinen kognitiven Fähigkeit ebenso als starker Prädiktor der Mathematikleistung ermittelt werden, jedoch zeigten sich keine geschlechtsspezifischen Unterschiede weder in der Mathematikleistung noch in der Selbstwirksamkeitserwartung. Die Autoren führen dies unter anderem auf den relativ geringen Stichprobenumfang ($N = 329$) und auf das Alter der Testpersonen (SchülerInnen) zurück, die zum Großteil erst die 9. oder 10. Klasse ($N = 206$) besuchten.

Bereits im Kapitel zu Stereotype Threat wurde dargelegt, dass die geschlechtsspezifischen Leistungsunterschiede in mathematischen, naturwissenschaftlichen und technischen Bereichen zuungunsten der Frauen ab der Sekundarstufe zu beobachten sind und immer deutlicher bzw. stärker werden (Hyde, Fennema & Lamon, 1990, zitiert nach Steele, 1997, S. 615).

Geschlechtsspezifische Unterschiede in der Selbstwirksamkeitserwartung könnten gemäß Pajares und Kranzler (1995) ebenso erst später relevant werden. Ein weiterer Erklärungsansatz der Autoren richtete sich auf die Generalität bzw. Spezifität der Erfassung des Konstrukts, demnach könnten zwar Unterschiede in der Selbstwirksamkeitserwartung zuungunsten der weiblichen Schüler bestehen, jedoch in Bezug auf den spezifischen Kontext, wie hier, weniger stark hervortreten als womöglich im allgemeinen.

Bei Hinz, Schumacher, Albani, Schmid und Brähler (2006) zeigte sich in Deutschland im Rahmen einer bevölkerungsrepräsentativen Untersuchung (N = 2019; 1070 Frauen und 949 Männer) der *Skala zur Allgemeinen Selbstwirksamkeitserwartung* (Jerusalem & Schwarzer, 1999), dass sich Männer eine höhere Allgemeine Selbstwirksamkeitserwartung zuschreiben als Frauen. Die Altersspanne der UntersuchungsteilnehmerInnen umfasste 16 bis 95 Jahre ($M = 48.5$, $SD = 17.4$) und es wurde festgestellt, dass sich jüngere Personen selbstwirksamer einschätzen als ältere. Demgemäß kann eine Verringerung in der Selbstwirksamkeitserwartung mit zunehmendem Alter angenommen werden.

Eine Untersuchung von Schwarzer et al. (1997) mit einer Stichprobe von insgesamt 1682 StudentInnen aus Deutschland (N = 430), Hongkong (N = 293) und Costa Rica (N = 959) verdeutlicht geschlechtsspezifische Unterschiede hinsichtlich der Allgemeinen Selbstwirksamkeitserwartung zugunsten der männlichen Testpersonen in der deutschen und in der chinesischen Stichprobe, während sich bei den Testpersonen aus Costa Rica keine Unterschiede zeigten.

Schwarzer, Mueller und Greenglass (1999) führten eine innovative Erhebungsmethode zur Allgemeinen Selbstwirksamkeitserwartung durch: Die Erfassung des Konstrukts im Rahmen einer Internetbefragung. An der Onlinestudie nahmen insgesamt 1437 Personen (583 Frauen und 762 Männer; N = 1345 gültige Fälle) teil, wobei die überwiegende Mehrheit (78 %) aus Amerika stammte und die Hälfte der TeilnehmerInnen unter 26 Jahren war. Die AutorInnen stellten fest, dass UntersuchungsteilnehmerInnen mit einer hohen Selbstwirksamkeitserwartung höhere akademische Leistungen aufweisen als Personen mit einer niedrigen Selbstwirksamkeitserwartung. Kein Zusammenhang konnte hingegen zwischen dem Lebensalter und der Selbstwirksamkeitserwartung beobachtet werden.

Die Untersuchungen zur Allgemeinen Selbstwirksamkeitserwartung deuten darauf hin, dass die Allgemeine Selbstwirksamkeitserwartung als universelles Konstrukt angenommen werden kann (Luszczynska et al., 2005; Schwarzer et al., 1997, 1999).

Bezugnehmend auf geschlechtsspezifische Unterschiede zeigte sich, dass sich die männlichen Testteilnehmer selbstwirksamer einschätzen als die weiblichen. Die Ergebnisse weisen darauf hin, dass Männer ein stärkeres allgemeines Vertrauen in ihre Fähigkeiten bekunden, stressreiche Situationen erfolgreich meistern zu können als Frauen. Ein Erklärungsansatz liegt womöglich in den Sozialisationsunterschieden zwischen Frauen und Männern (Schwarzer et al., 1999).

Im Rahmen einer Längsschnittuntersuchung des Modellversuchs „Selbstwirksame Schulen“ in Deutschland wurde der Einfluss der schulischen Selbstwirksamkeitserwartung unter anderem auf die Leistung (Mathematik, Deutsch und Englisch), Lernfreude und das Wohlbefinden der SchülerInnen (N = 3072 SchülerInnen der siebten bis zwölften Jahrgangsstufen) analysiert (Edelstein, Hopf, Jerusalem & Schwarzer, 1997; Jerusalem & Mittag, 1999). Das primäre Ziel dieses Projekts stellte die Förderung von Selbstwirksamkeitserwartungen in den zehn teilnehmenden Pilotschulen dar, wobei kein allgemeines Programm durchgeführt wurde, sondern die Schulen eigene Schwerpunkte setzten (Hofmann, 1997; Mittag, Kleine & Jerusalem, 2002).

Die Ergebnisse der ersten Erhebung zeigten, dass SchülerInnen mit einer hohen Selbstwirksamkeitserwartung unabhängig von Geschlecht und Jahrgangsstufe bessere Noten in den drei oben genannten Fächern aufwiesen als SchülerInnen mit einer niedrigen Selbstwirksamkeitserwartung. Des Weiteren stellten die Autoren fest, dass SchülerInnen mit einer hohen Selbstwirksamkeitserwartung eine höhere Lernfreude in den drei genannten Fächern angaben als SchülerInnen mit einer niedrigen Selbstwirksamkeitserwartung. Den Ergebnissen zufolge kann eine Verringerung der Lernfreude über die Jahrgangsstufen beobachtet werden und es zeigen sich geschlechtsspezifische Unterschiede: Mädchen geben eine niedrigere Lernfreude in Mathematik an und Burschen berichten über eine niedrigere Lernfreude in Deutsch und Englisch. Darüber hinaus wurde ein moderierender Effekt der Selbstwirksamkeitserwartung auf den Zusammenhang zwischen der Mathematik-Leistung und der Lernfreude in Mathematik festgestellt: Bei SchülerInnen mit einer hohen Selbstwirksamkeitserwartung zeigt sich im Gegensatz zu SchülerInnen mit einer niedrigen Selbstwirksamkeitserwartung, dass die Lernfreude für Mathematik angesichts *sehr guter* Mathematikleistungen zunimmt und die Lernfreude für Mathematik angesichts *schlechter* Mathematikleistungen nicht abnimmt (Edelstein et al., 1997; Jerusalem & Mittag, 1999).

Bezugnehmend auf das Wohlbefinden gaben SchülerInnen mit einer hohen Selbstwirksamkeitserwartung ein deutlich höheres Maß an optimistischer Lebenseinstellung sowie Zufriedenheit mit der Schule und ein deutlich geringeres Maß an Depressivität und Hilflosigkeit an im Vergleich zu SchülerInnen mit einer niedrigen Selbstwirksamkeitserwartung (Jerusalem & Mittag, 1999).

Die Evaluation dieses dreijährigen Modellversuchs „Selbstwirksame Schulen“ mit insgesamt vier Messzeitpunkten basiert auf einer Teilstichprobe (N = 434 SchülerInnen, die beim 1. Messzeitpunkt die siebte oder achte Klasse besuchten). Hinsichtlich der Entwicklung der schulbezogenen Selbstwirksamkeitserwartung zeigte sich, dass von 427 SchülerInnen bei 45 % eine Erhöhung, bei 27,6 % ein gleich bleibendes Niveau und bei 27,4 % eine Verminderung der schulbezogenen Selbstwirksamkeitserwartung angenommen werden kann. Gemäß den Ergebnissen ist die Entwicklung der Selbstwirksamkeitserwartung stärker verbunden mit der Entwicklung der Lernfreude in Mathematik sowie mit Unterschieden in der Mathematikleistung im Vergleich zu den Fächern Deutsch und Englisch. Ebenso zeigte sich ein stärkerer Zusammenhang zwischen der Lernfreude für Mathematik und der Leistung in Mathematik im Vergleich zu Deutsch und Englisch. Ein möglicher Erklärungsansatz stellt laut Autoren die stärkere Transparenz der Leistungsrückmeldung im Fach Mathematik dar. Darüber hinaus weisen die Evaluationsergebnisse darauf hin, dass Selbstwirksamkeitserwartungen stärker die schulischen Leistungen beeinflussen als umgekehrt (Mittag et al., 2002).

Bei Lent, Brown und Larkin (1984) konnte bei einer Untersuchung von 42 StudentInnen, die eine Karriere im naturwissenschaftlichen oder technischen Bereich anstreben, ebenso ein positiver Zusammenhang zwischen der Selbstwirksamkeitserwartung und ihren bisherigen schulischen Leistungen, die zum Vergleich herangezogen wurden, festgestellt werden. Die Untersuchung, die im Rahmen eines Karriereplanungsseminars durchgeführt wurde, zeigte einen positiven Zusammenhang zwischen der Selbstwirksamkeitserwartung und den akademischen Leistungen, die sich auf das weitere Studienjahr (kumulierte durchschnittliche Leistung insgesamt) und das darauf folgende Studienjahr (durchschnittliche Leistung nur in den naturwissenschaftlichen oder technischen Bereichen) bezogen. Eine bedeutende Erkenntnis der Studie, stellt zudem der Befund dar, dass die Selbstwirksamkeitserwartung mit dem Verbleib in den angestrebten Studiengängen positiv zusammenhängt.

In einer weiteren Studie der Autoren mit einem vergleichbaren Design konnte ebenso ein positiver Zusammenhang zwischen der Selbstwirksamkeitserwartung und den akademischen Leistungen sowie dem Verbleib im Studium belegt werden. Darüber hinaus kann die Selbstwirksamkeitserwartung als Prädiktor dieser beiden Variablen und der wahrgenommenen beruflichen Karrieremöglichkeiten angenommen werden (Lent, Brown & Larkin, 1986).

4.5 Selbstwirksamkeitserwartung und Stereotype Threat

Der positive Zusammenhang zwischen der Selbstwirksamkeitserwartung und Leistung (z.B.: Multon et al., 1991) konnte im vorigen Abschnitt verdeutlicht werden. Bezugnehmend auf das Phänomen Stereotype Threat, erfolgte im Unterpunkt 3.4 die Darlegung etwaiger Mechanismen, die die Leistungsminderung unter Stereotype Threat bewirken könnten. In diesem Sinne wird auch die Selbstwirksamkeitserwartung als ein möglicher vermittelnder Mechanismus der Stereotype Threat-Effekte vermutet (z.B.: Spencer et al., 1999).

Spencer et al. (1999) untersuchten die Selbstwirksamkeitserwartung als etwaigen Mediator der Leistungsminderung von Frauen unter Stereotype Threat (zur erhobenen Variable Ängstlichkeit siehe S. 13). Die Ergebnisse zeigten, dass kein Unterschied in der Selbstwirksamkeitserwartung zwischen den Frauen in der *Stereotype Threat-Bedingung* und den Frauen in der *Non-Threat-Bedingung* beobachtet werden kann. Darüber hinaus konnte kein Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Mathematikleistung festgestellt werden. Demgemäß stellt die Selbstwirksamkeitserwartung keine Mediatorvariable der in der Studie demonstrierten Stereotype Threat-Effekte dar. Hinsichtlich des nicht gefundenen mediiierenden Effekts der Selbstwirksamkeitserwartung könnte jedoch auch dem relativ geringen Stichprobenumfang eine ursächliche Rolle zukommen.

Bei Oswald und Harvey (2000) wurden Studentinnen 50 mathematische Problemlöseaufgaben mit einer Bearbeitungszeit von insgesamt 30 Minuten vorgegeben. Mittels Gedankenauflistung (Anführung aller Gedanken, die den Untersuchungsteilnehmerinnen während der Testbearbeitung in den Sinn gekommen sind) erfolgte die Erfassung der Selbstwirksamkeitserwartung, wobei die Interrater-Reliabilität für die Selbstwirksamkeitserwartung 93 Prozent betrug. Die AutorInnen kamen zu dem Ergebnis, dass Studentinnen in der *Stereotype Threat-Bedingung* eine geringere Mathematikleistung als ihre Kolleginnen in der *Non-Threat-Bedingung* erzielten, jedoch konnte auch in dieser Studie die Selbstwirksamkeitserwartung nicht als Mediator des Stereotype Threat-Effekts eruiert werden.

Gemäß Smith (2004, zitiert nach Chung, Ehrhart, Holcombe Ehrhart, Hattrup & Solamon, 2010, S. 88) scheint es überraschend, dass der erwartete mediiierende Effekt

der Selbstwirksamkeitserwartung auf die Leistungsminderung unter Stereotype Threat in den durchgeführten Studien kaum untermauert werden kann.

In diesem Zusammenhang untersuchten Chung et al. (2010) ein Modell, in dem nicht nur die Selbstwirksamkeitserwartung, sondern auch die Ängstlichkeit integriert wurde. Die Erhebung erfolgte im Rahmen eines Berufsauswahlverfahrens mit insgesamt 150 BewerberInnen (57,2 % afroamerikanische und 41,4 % weiße BewerberInnen). Es zeigte sich, dass die weißen BewerberInnen eine höhere Leistung in dem vorgegebenen Berufswissenstest erzielten als ihre afroamerikanischen KollegInnen. Übereinstimmend mit den theoretischen Modellannahmen konnte festgestellt werden, dass ein positiver Zusammenhang zwischen der Wahrnehmung von Stereotype Threat und der Ängstlichkeit sowie ein negativer Zusammenhang zwischen der Ängstlichkeit und der Selbstwirksamkeitserwartung beobachtet werden kann. Die Selbstwirksamkeitserwartung wiederum übte einen positiven Effekt auf die Leistung im Berufswissenstest aus. Darüber hinaus konnten in der Untersuchung die Ängstlichkeit und die Selbstwirksamkeitserwartung als Mediatoren des Stereotype Threat-Effekts ermittelt werden. Demgemäß könnten die Ängstlichkeit und die Selbstwirksamkeitserwartung als Mechanismen der Leistungsminderung unter Stereotype Threat angenommen werden.

5. Risikoverhalten

Unter dem Begriff *Risikoverhalten* versteht man nach Trimpop (1994) „any consciously, or non-consciously controlled behavior with a perceived uncertainty about its outcome, and/or about its possible benefits or costs for the physical, economic or psycho-social well-being of oneself or others” (S. 9).

Im Rahmen dieser allgemeinen Definition betont der Autor zum einen die Unsicherheit in Bezug auf das Ergebnis und zum anderen die etwaig verbundenen Konsequenzen (Trimpop, 1994).

Risikoverhalten ist gemäß Byrnes, Miller und Schafer (1999) „adaptive or maladaptive. It is maladaptive whenever the benefits of some activity are far less likely to occur than the potential hazards. It is adaptive whenever the converse is true” (S. 368).

Nachfolgend werden geschlechtsspezifische Unterschiede in der Risikoverhaltens-Tendenz thematisiert.

Cárdenas, Dreber, von Essen und Ranehill (2012) demonstrierten bei einer Stichprobe von 1240 SchülerInnen aus Schweden und Kolumbien geschlechtsspezifische Unterschiede im Risikoverhalten. In der Untersuchung mussten die SchülerInnen ($M = 10.90$ Jahre, $SD = 0.91$, $N = 1120$ gültige Fälle) Entscheidungen treffen, wobei sie jeweils entweder einen Münzwurf, der null respektive zehn Punkte brachte (Risiko) oder eine sichere Option (sukzessiver Zugewinn von einer geringeren Punktezahl) wählen konnten. Die AutorInnen stellten fest, dass die Burschen, unabhängig von der Nation, eine höhere Risikoverhaltens-Tendenz im Vergleich zu den Mädchen aufwiesen. Die geschlechtsspezifischen Unterschiede im Risikoverhalten waren in Kolumbien deutlich stärker ausgeprägt als in Schweden.

Geschlechtsspezifische Unterschiede im Risikoverhalten scheinen auch bei finanziellen Entscheidungen zu bestehen, wie Charness und Gneezy (2012) feststellten. Bei einer Überprüfung von 15 Studien zu Investmententscheidungen, zeigte sich, dass bei Männern eine höhere Risikoverhaltens-Tendenz angenommen werden kann im Vergleich zu Frauen.

Eine Metaanalyse von Byrnes et al. (1999), basierend auf insgesamt 150 Untersuchungen mit über 100.000 Testpersonen, weist auf eine höhere Risikoverhaltens-Tendenz zugunsten der männlichen Untersuchungsteilnehmer in der

überwiegenden Mehrheit der untersuchten 16 Risikoverhaltensbereiche (z.B.: Raten, Autofahren, Glücksspiel) hin. Die jeweiligen Untersuchungen wurden gemäß des verwendeten Aufgabentyps einer von drei Kategorien zugeordnet, hierbei umfasste die Mehrzahl der Untersuchungen die Kategorie „bekundetes Risikoverhalten“ gefolgt von den Kategorien „beobachtetes Risikoverhalten“ und „hypothetisches Risikoverhalten“ (z.B.: Wahl zwischen potentiellen Optionen). Es zeigte sich, dass geschlechtsspezifische Unterschiede im Risikoverhalten deutlicher bei beobachtetem Risikoverhalten angenommen werden können im Vergleich zu bekundetem Risikoverhalten (z.B.: Angabe der Häufigkeit bestimmter Verhaltensweisen). Weiters wiesen die Ergebnisse darauf hin, dass der Zuwachs der geschlechtsspezifischen Unterschiede mit zunehmendem Alter der Testpersonen vermutlich geringer wird; die geschlechtsspezifischen Unterschiede in der Alterskategorie ≥ 22 Jahre waren niedriger ausgeprägt als bei SchülerInnen in der Alterskategorie von 14 – 17 Jahren. Die Autoren stellten fest, dass geschlechtsspezifische Unterschiede in der Risikoverhaltens-Tendenz in Abhängigkeit vom Risikoverhaltensbereich und der Alterskategorie unterschiedlich stark angenommen werden können. Bei sieben Risikoverhaltensbereichen konnte eine durchschnittliche Effektgröße von $d > 0.20$ ermittelt werden. Insgesamt kann ein geschlechtsspezifischer Unterschied in der Risikoverhaltensrate von sechs Prozent beobachtet werden.

In Bezug auf den Leistungskontext demonstrierten Ben-Shakhar und Sinai (1991) eine höhere Risikoverhaltens-Tendenz zugunsten der männlichen Untersuchungsteilnehmer bei zwei Multiple-Choice Tests. In der Untersuchung wurde SchülerInnen der 9. Schulstufe und StudienbewerberInnen jeweils eine Testbatterie vorgegeben, wobei die StudentInnen instruiert wurden, dass sie bei den Multiple-Choice Items raten sollen, wenn sie sich nicht ganz sicher sind, wohingegen die SchülerInnen nicht darauf hingewiesen wurden. Es zeigte sich, dass die männlichen Testteilnehmer in den meisten Untertests eine höhere Leistung im Vergleich zu ihren weiblichen Kolleginnen erzielten. Darüber hinaus stellten die Autoren eine höhere Tendenz der weiblichen Untersuchungsteilnehmer fest, Items auszulassen. An dieser Stelle ist darauf hinzuweisen, dass sich dies auch bei Untertests, die keine geschlechtsspezifischen Leistungsunterschiede zugunsten der männlichen Testpersonen verdeutlichen, zeigte. Hinsichtlich der Instruktion zum Raten konnte zwar bei den Studentinnen eine etwas geringere Tendenz Items auszulassen im Vergleich zu den Schülerinnen festgestellt werden, jedoch blieben die geschlechtsspezifischen Unterschiede in der Ratetendenz zugunsten der männlichen Testteilnehmer bestehen.

Bei Hanna (1986) zeigte sich sogar eine deutlich höhere Tendenz von weiblichen Schülern, Items auszulassen im Vergleich zu ihren männlichen Kollegen: Ein Mathematiktest bestehend aus Multiple-Choice Items mit jeweils fünf Antwortmöglichkeiten und immer nur einer richtigen Lösung wurde SchülerInnen mittels Prä- und Posttest vorgegeben. Bei der Studie, an der insgesamt 3523 SchülerInnen (50,3 % Burschen und 49,7 % Mädchen) der 8. Schulstufe in Kanada teilnahmen zeigte sich sowohl beim Prä- als auch beim Posttest, dass die Burschen in den Untertests Maßzahlen und Geometrie (zwei von insgesamt fünf mathematikbezogenen Untertests) eine höhere Leistung im Vergleich zu ihren weiblichen Schulkolleginnen erzielten. Darüber hinaus weisen die Untersuchungsergebnisse auf eine sehr geringe Risikoverhaltens-Tendenz der Schülerinnen hin. Es konnte eine deutlich erhöhte Auslassungsrate von den weiblichen Schülern (bei 116 Items) im Vergleich zu ihren männlichen Schulkollegen (bei 17 Items) beobachtet werden.

Von Schrader und Ansley (2006) stellten hingegen nur geringe Unterschiede in der Tendenz Items auszulassen zwischen den männlichen und weiblichen SchülerInnen fest. Die AutorInnen analysierten bei SchülerInnen der dritten, siebten und elften Schulstufe im US-Bundesstaat Iowa die erreichten Punkte bei zwei Testverfahren im Zeitraum von 1980 bis 2000. Im Rahmen der zwanzigjährigen Zeitspanne konnte beobachtet werden, dass die Tendenz Items auszulassen sich verringerte, sobald die Leistungen anstiegen. In der dritten Schulstufe zeigte sich eine höhere Auslassungstendenz bei mathematischen Untertests zugunsten der weiblichen Schüler und bei Untertests zu Vokabeln und Lesen zugunsten der männlichen Schüler. Insgesamt weisen die Untersuchungsergebnisse auf keinen bedeutenden Einfluss von geschlechtsspezifischen Unterschieden in der Auslassungstendenz auf die Testleistung hin.

Die Ergebnisse von Bolger und Kellaghan (1990) könnten jedoch mit unterschiedlichen Risikoverhaltens-Tendenzen der Geschlechter in Zusammenhang stehen. Die Autoren analysierten etwaige geschlechtsspezifische Leistungsunterschiede in Mathematik, Irisch und Englisch bei der Vorgabe der jeweiligen Tests einerseits mittels Multiple-Choice Items (es gab keine „Rate-Instruktion“) und andererseits anhand freiem Antwortformat. In der Untersuchung wurde eine Stichprobe von 1497 SchülerInnen aus Irland (739 Burschen und 758 Mädchen) mit einem Durchschnittsalter von etwa 15 Jahren herangezogen. Bei den Tests, die Mitte und Ende des Schuljahres vorgegeben wurden, zeigte sich, dass Burschen bei dem Mathematiktest sowohl mittels Multiple-

Choice Items als auch bei freiem Antwortformat eine höhere Leistung als ihre weiblichen Kolleginnen erzielten. Die Mädchen erreichten bei den Irishtests (freies Antwortformat und MC) und bei dem Englishtest (freies Antwortformat) eine bessere Leistung als die Burschen, allerdings war der Leistungsunterschied bei dem Multiple-Choice Englishtest zugunsten der männlichen Schüler sehr gering. Weitere Analysen weisen darauf hin, dass Items mit freiem Antwortformat die Leistung der weiblichen Schüler begünstigen, während dies bei den männlichen Schülern bei Multiple-Choice Items angenommen werden kann. Ein Erklärungsansatz für die geringere Leistung der Mädchen bei Multiple-Choice Items könnte die höhere Rate- bzw. Risikoverhaltens-Tendenz ihrer männlichen Schulkollegen darstellen.

Zu einem vergleichbaren Ergebnis gelangten auch Arthur und Everaert (2012): In ihrer Untersuchung zeigte sich im akademischen Kontext bei einer Rechnungswesen-Prüfung, dass Multiple-Choice Items die Leistung der männlichen Studenten begünstigen, wohingegen „constructed response“-Fragen bei den weiblichen Studentinnen leistungsfördernd zu wirken scheinen. Dieser Befund konnte auch bei StudienabsolventInnen ($M \sim 23$ Jahre) beobachtet werden und verdeutlicht die Robustheit dieses Effekts. Gemäß AutorInnen stellt eine höhere Rate-Tendenz der männlichen Testteilnehmer einen möglichen Faktor für die geschlechtsspezifischen Unterschiede dar, jedoch scheint dies die Ergebnisse nicht vollständig erklären zu können.

6. Zielsetzungen, Fragestellungen und Hypothesen

6.1 Zielsetzungen und Fragestellungen

Die vorliegende Arbeit widmet sich zunächst der Untersuchung eines möglichen Zusammenhangs zwischen der Selbstwirksamkeitserwartung und den Leistungen in EMS-ähnlichen Untertests. Weiters wird analysiert, ob es Unterschiede in der Selbstwirksamkeitserwartung in Abhängigkeit vom Geschlecht gibt. Darüber hinaus soll geklärt werden, ob das Geschlecht, Stereotype Threat und Risikoverhalten die Leistungen in EMS-ähnlichen Untertests beeinflussen. Ein vorrangiges Ziel stellt hierbei die Überprüfung eines etwaigen Einflusses der Selbstwirksamkeitserwartung als Kovariate dar.

6.2 Hypothesen

Selbstwirksamkeitserwartung (SWE)

H₁(1): Es gibt einen signifikanten Zusammenhang zwischen der Selbstwirksamkeitserwartung und den Leistungen in EMS-ähnlichen Untertests.

H₁(2): Es gibt signifikante Unterschiede und signifikante Wechselwirkungen in der Selbstwirksamkeitserwartung in Abhängigkeit der drei Faktoren: zwischen Männern und Frauen, zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) und zwischen den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

H₁(3): Es gibt einen signifikanten Unterschied in der Selbstwirksamkeitserwartung zwischen der Intention, Medizin zu studieren (ja/nein).

Leistungen in EMS-ähnlichen Untertests

H₁(4): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen.

H₁(5): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung).

H₁(6): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

H₁(7): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung).

H₁(8): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

H₁(9): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

H₁(10): Es gibt eine signifikante Interaktion höherer Ordnung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

Leistungen in EMS-ähnlichen Untertests unter Berücksichtigung der SWE als Kovariate

H₁(11): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(12): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(13): Es gibt einen signifikanten Unterschied in den Leistungen bei EMS-ähnlichen Untertests zwischen den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(14): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(15): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(16): Es gibt eine signifikante Wechselwirkung in den Leistungen bei EMS-ähnlichen Untertests zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

H₁(17): Es gibt eine signifikante Interaktion höherer Ordnung in den Leistungen bei EMS-ähnlichen Untertests zwischen Männern und Frauen sowie den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) sowie den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate.

Risikoverhaltens-Tendenz

H₁(18): Es gibt einen signifikanten Unterschied in der Risikoverhaltens-Tendenz zwischen den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten).

H₁(19): Es gibt einen signifikanten Unterschied in der Risikoverhaltens-Tendenz zwischen Männern und Frauen.

Empirischer Teil

7. Methode

Nachfolgend wird der Untersuchungshintergrund der durchgeführten Studie dargestellt. Daran anschließend erfolgt die Darlegung des Untersuchungsplans sowie der intendierten Stichprobe und der verwendeten Erhebungsinstrumente.

Die wissenschaftliche Untersuchung dieser Diplomarbeit ist Teil einer umfangreichen Studie, die von der Medizinischen Universität Wien im Jahr 2009 zur Klärung von geschlechtsspezifischen Unterschieden beim EMS (Eignungstest für das Medizinstudium), in Auftrag gegeben wurde.

An der Durchführung der Studie, unter der Leitung von Frau Ao. Univ. Prof. Dr. Mag. Ulrike Willinger, waren neben mir vier weitere Diplomandinnen der Fakultät für Psychologie involviert: Katrin Anzirk, Silvia Hameseder, Gabriele Hangl und Helene Lagger. Jede Diplomandin widmet sich in ihrer Arbeit einem spezifischen Schwerpunkt³ der Untersuchung.

7.1 Untersuchungsplan und intendierte Stichprobe

Die Intention der Studie stellte im Rahmen eines multivariaten Designs, die Erforschung von geschlechtsspezifischen Unterschieden in den Leistungen von EMS-ähnlichen Untertests und Theory of Mind – Aufgaben, dar. Hierbei sollte der Einfluss von Stereotype Threat und Risikoverhalten (*soziale Faktoren*) sowie des Hormonstatus vor und nach der Testung (*biologischer Faktor*) analysiert werden. Ein weiterer wichtiger Punkt der Studie lag in der Untersuchung möglicher *Mediatorvariablen* (Ängstlichkeit, Selbstwirksamkeit und Leistungsmotivation).

Im Hinblick auf die intendierte Stichprobe, sollten Erhebungen an 1200 SchülerInnen der 12. Schulstufe von allgemeinbildenden höheren Schulen in Wien durchgeführt werden. Anhand von Tabelle 2 ist das geplante Design der Studie ersichtlich. Die

³ Die Darlegung der Ergebnisse zu den weiteren untersuchten Aspekten ist den Arbeiten meiner vier Kolleginnen zu entnehmen: Theory of Mind (Katrin Anzirk), schlussfolgerndes Denken (Silvia Hameseder), Leistungsmotivation (Gabriele Hangl) und Merkfähigkeit (Helene Lagger).

Zuweisung der Testpersonen zu den beiden experimentellen Faktoren (vier *Stereotype Threat-Bedingungen* und drei *Risikoverhalten-Bedingungen*) und dem festen Faktor *Geschlecht*, somit auf die 24 (4*3*2) Faktorstufenkombinationen, sollte randomisiert erfolgen.

Tabelle 2: geplantes Studiendesign

Variablen	Versuchsbedingungen															
	A			B			C			D						
Hormonmessung																
Stereotype Threat	A			B			C			D						
Risikoverhalten	1	2	3	1	2	3	1	2	3	1	2	3				
Geschlecht	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂
EMS-ähnlicher Test																
Theory of Mind-Aufgaben																
Hormonmessung																
Ängstlichkeit																
Selbstwirksamkeit																
Leistungsmotivation																

Nachfolgend werden die geplanten Untersuchungsbedingungen der beiden experimentellen Faktoren (Stereotype Threat und Risikoverhalten) dargelegt.

Stereotype Threat – Untersuchungsbedingungen:

A) Aktivierung

In dieser Bedingung wird den SchülerInnen mittels der Instruktion mitgeteilt, dass Frauen bessere Leistungen im Gedächtnis aufweisen, während Männer bessere Leistungen in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen, im Verstehen von komplexen Texten und im naturwissenschaftlichen Wissen erzielen (vgl. Smith & White, 2002).

B) Neutrale Bedingung

Die UntersuchungsteilnehmerInnen werden darauf hingewiesen, dass es keine geschlechtsspezifischen Unterschiede in den Leistungen zwischen Männern und Frauen gibt (vgl. Spencer et al., 1999).

C) Aufklärung

Die Testpersonen erhalten Informationen zum Phänomen Stereotype Threat. Den SchülerInnen wird mitgeteilt, dass eine möglicherweise auftretende Ängstlichkeit während der Testbearbeitung auch ein Resultat negativer Stereotype sein könnte, die in der Gesellschaft weit verbreitet sind und, dass die Ängstlichkeit mit den gegenwärtigen Fähigkeiten im Test gut abzuschneiden nichts zu tun hat (vgl. Johns et al., 2005).

D) Negativ-Aktivierung

Den UntersuchungsteilnehmerInnen wird vermittelt, dass Männer bessere Leistungen im Gedächtnis erzielen, während Frauen bessere Leistungen in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen, im Verstehen von komplexen Texten und im naturwissenschaftlichen Wissen aufweisen.

Risikoverhalten– Untersuchungsbedingungen:

1) Raten

Den Testpersonen wird mitgeteilt, dass sie bei den nachfolgenden Aufgaben versuchen sollen zu raten, wenn sie die richtige Lösung nicht wissen (vgl. Ben-Shakhar & Sinai, 1991).

2) Ohne Instruktion

Die Schülerinnen werden nur darauf hingewiesen, dass sie die nachfolgenden Aufgaben aufmerksam und konzentriert durcharbeiten sollen.

3) „Sicheres“ Antworten

Den Testpersonen werden instruiert, dass sie bei den nachfolgenden Aufgaben versuchen sollen nur dann eine Antwort anzukreuzen, wenn sie sich ganz sicher sind.

Die jeweiligen Instruktionen (Kombination aus einer Stereotype Threat-Bedingung und einer Risikoverhalten-Bedingung) sollten den UntersuchungsteilnehmerInnen vor der Bearbeitung jedes einzelnen der insgesamt sieben EMS-ähnlichen Untertests (Quantitative und formale Probleme, Schlauchfiguren, Textverständnis, Medizinisch-naturwissenschaftliches Grundverständnis, Diagramme und Tabellen, Figuren lernen

und Fakten lernen) und vor den Theory of Mind – Aufgaben in Erinnerung gerufen werden, indem auf Zwischenblättern die Instruktionvorgabe nochmals erfolgen sollte. Anhand einer Vorstudie sollte die Wirksamkeit der Stereotype Threat-Bedingungen und Risikoverhalten-Bedingungen überprüft werden. Die Untersuchung der 1200 SchülerInnen war in Gruppentestungen vorgesehen und sollte insgesamt zwei Unterrichtsstunden (100 Minuten) dauern. Mittels Speichelproben vor der Bearbeitung des EMS-ähnlichen Test und nach den Theory of Mind – Aufgaben sollten bei den SchülerInnen folgende Hormone bestimmt werden: Cortisol, DHEA, Testosteron, Dopamin, Chromogranin A, Östradiol und Progesteron. Die gesamte Planung und Durchführung der Studie sowie die Dateneingabe und die Auswertung der Ergebnisse sollten im Zeitraum von Oktober 2009 bis Juni 2010 erfolgen.

7.2 Erhebungsinstrumente

In diesem Abschnitt erfolgt die Darlegung der Erhebungsverfahren gemäß der Reihenfolge, in der sie in der Studie angewendet wurden (siehe Tabelle 3).

Tabelle 3: Erhebungsinstrumente

Erhebungsinstrumente	
Soziodemographischer Fragebogen	
Wissenstest	basierend auf Halpern und Tan (2001)
EMS-ähnliche Untertests	(ZTD, 2009) (ITB Consulting, 2008)
Schlauchfiguren	(Stumpf & Fay, 1983)
Theory of Mind – Aufgaben	(Willinger, Schmöger, Müller & Auff, in Vorbereitung)
Leistungsmotivationsinventar	(Schuler & Prochaska, 2001)
Skala zur Allgemeinen Selbstwirksamkeitserwartung	(Jerusalem & Schwarzer, 1999)
Game of Dice Task	(Brand et al., 2005)

7.2.1 Soziodemographischer Fragebogen

Mittels des soziodemographischen Fragebogens wurden zunächst folgende Merkmale der SchülerInnen erhoben: Geschlecht, Alter, Staatsbürgerschaft, besuchter Ausbildungszweig in der Oberstufe, Maturafächer und Wiederholung von zumindest

einer Schulstufe. Weiters erfolgte die Erfassung der höchsten abgeschlossenen Ausbildung der Eltern und deren Beruf sowie die Freizeitaktivitäten der UntersuchungsteilnehmerInnen. Im letzten Teil des Fragebogens wurden Berufswunsch der Testpersonen, geplante Ausbildung, Intention ein Medizinstudium zu beginnen sowie die angestrebte Tätigkeitsrichtung nach Absolvierung eines Medizinstudiums und die bisherige Befassung mit dem EMS erhoben.

7.2.2 Wissenstest

Der *Wissenstest* wurde auf Basis des „Beliefs About the Cognitive Abilities of Females and Males Questionnaire“ von Halpern und Tan (2001) konstruiert und besteht aus insgesamt 14 Items. Anhand des Wissenstest werden einerseits die persönlichen und andererseits die geschlechtsspezifischen Fähigkeitszuschreibungen erfasst und die Wirksamkeit der Stereotype-Threat Untersuchungsbedingungen überprüft. Der erste Teil des Wissenstest enthält sieben Items zur geschlechtsspezifischen Fähigkeitszuschreibung, wobei sechs Aufgaben den in weiterer Folge vorgegeben EMS-ähnlichen Untertests zuzuordnen sind und ein Item der Theory of Mind. Die UntersuchungsteilnehmerInnen sollten jede Aufgabe hinsichtlich der Wahrscheinlichkeit (in Prozentwerten) einschätzen, dass es sich bei der Lösung des jeweiligen Items respektive der Fähigkeit um einen Mann bzw. eine Frau handle und die Summe der jeweiligen Prozentwerte sollte immer 100 % ergeben. Die Abbildung 2 stellt ein Itembeispiel zur geschlechtsspezifischen Fähigkeitszuschreibung dar.

4) Denken Sie an eine Person, die Sie nicht kennen. Der Person wurde ein Text zu einem medizinisch-naturwissenschaftlichen Thema (z.B. Herzkreislaufsystem) vorgelegt zu dem sie anschließend Verständnisfragen beantworten musste.

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

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

Abbildung 2: Itembeispiel geschlechtsspezifische Fähigkeitszuschreibung; Untertest Medizinisch-naturwissenschaftliches Grundverständnis (Hänsgen & Spicher, 2010)

Darauffolgend wurde zu den sechs Untertests und zur Theory of Mind jeweils ein Item vorgegeben, bei dem die UntersuchungsteilnehmerInnen beurteilen sollten, ob sie über die angeführte Fähigkeit verfügen (siehe Abbildung 3).

Die Itemanzahl und die zur Verfügung stehende Bearbeitungszeit in den jeweiligen EMS-ähnlichen Untertests, mit Ausnahme der Schlauchfiguren, wurden analog zum EMS so gewählt, dass es praktisch nicht möglich war, alle Items zu lösen und die Testung entsprechend dem EMS ähnlich hohe Anforderungen an die UntersuchungsteilnehmerInnen stellt (vgl. ZTD, 2009).

Nachfolgend werden die sechs Untertests dargestellt.

Quantitative und formale Probleme

Mittels des Untertests *Quantitative und formale Probleme* werden die quantitativen und formalen Kompetenzen bei medizinischen und naturwissenschaftlichen Fragestellungen erfasst (Hänsgen & Spicher, 2010).

Im Rahmen der Studie wurden 16 Items mit einer Bearbeitungszeit von 10 Minuten vorgegeben. Die UntersuchungsteilnehmerInnen sollten aus den Multiple-Choice-Aufgaben mit jeweils fünf Antwortmöglichkeiten die (einzig) richtige Lösung ankreuzen. Anhand der Abbildung 4 wird ein Itembeispiel zu diesem Untertest dargestellt.

3) Eine Broteinheit (BE) ist definiert als diejenige Nahrungsmenge in Gramm, die 12 g Kohlenhydrate enthält. Bei der Verbrennung von 1 g Kohlenhydraten im Organismus werden 16 Kilojoule (kJ) an Energie frei.
Ein Patient, der auf Diät gesetzt ist, soll pro Tag 4800 kJ zu sich nehmen, ein Fünftel davon in Kohlehydraten.

Wie viele BE sind dies täglich?

(A)	60	BE
(B)	25	BE
(C)	6	BE
(D)	5	BE
(E)	0,5	BE

Abbildung 4: Itembeispiel Quantitative und formale Probleme (Hänsgen & Spicher, 2010, S. 93)

Schlauchfiguren

Der Test *Schlauchfiguren* von Stumpf und Fay (1983) wurde zur Erhebung des räumlichen Vorstellungsvermögens eingesetzt. Dieser Test besteht aus 21 Items mit einer Bearbeitungszeit von 12 Minuten. Bei jeder Aufgabe werden zwei durchsichtige Würfel, die ein gedrehtes Kabel enthalten, abgebildet. Auf dem linken Bild wird immer die Vorderansicht des Würfels dargestellt und das rechte Bild zeigt den gleichen Würfel aus einem anderen Blickwinkel. Die Testpersonen sollten ankreuzen, aus welcher Perspektive (rechts, links, unten, oben oder hinten) der rechte Würfel zu sehen ist (Stumpf & Fay, 1983).

Nachfolgend wird ein Itembeispiel aus dem EMS-Untertest Schlauchfiguren angeführt (siehe Abbildung 5).

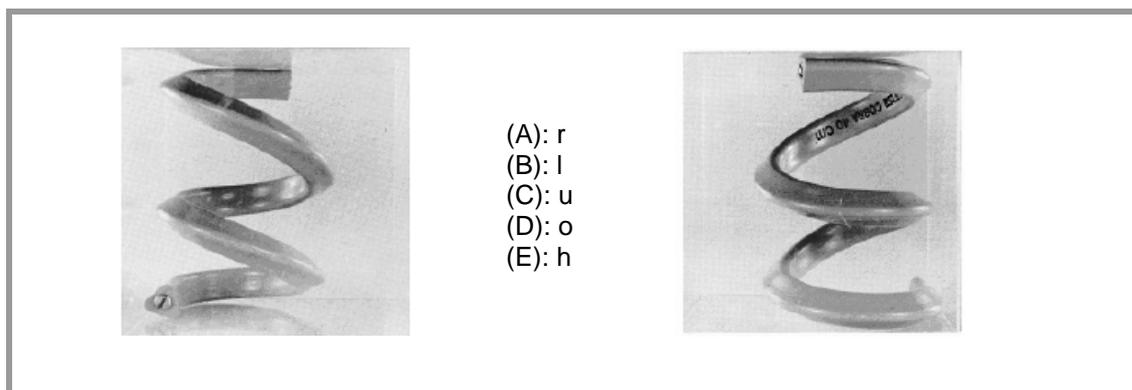


Abbildung 5: Itembeispiel Schlauchfiguren (Hänsgen & Spicher, 2010, S. 94)

Medizinisch-naturwissenschaftliches Grundverständnis

Anhand dieses Untertests wird sowohl die Verständnisfähigkeit als auch die Schlussfolgerungsfähigkeit bei medizinisch naturwissenschaftlichen Fragestellungen erfasst (Hänsgen & Spicher, 2010).

Die Abbildung 6 zeigt ein Itembeispiel zu diesem Untertest.

Im Rahmen der Studie erfolgte die Vorgabe von 16 Items mit einer Bearbeitungszeit von 10 Minuten

Im Kindesalter kann das Zentrum für Sprache, Spracherwerb und Sprachverständnis noch in der linken oder in der rechten Hälfte (Hemisphäre) des Gehirns in einem umschriebenen Hirnrindengebiet (sog. Sprachregion) angelegt werden. Spätestens im zwölften Lebensjahr sind die sprachlichen Fähigkeiten jedoch fest in einer der beiden Hemisphären verankert, und zwar bei den Rechtshändern in der Regel links, bei den Linkshändern in der Mehrzahl ebenfalls links, zum Teil aber auch rechts; die korrespondierende Region der Gegenseite hat zu diesem Zeitpunkt bereits andere Funktionen fest übernommen. Welche der nachfolgenden Aussagen lässt bzw. lassen sich aus diesen Informationen ableiten?

Bei irreversiblen Hirnrindenverletzungen im Bereich der sogenannten Sprachregion der linken Hemisphäre...

- I. kommt es bei erwachsenen Linkshändern in der Regel zu keinen wesentlichen Sprachstörungen.
 - II. kommt es bei einem Vorschulkind in der Regel zu einer bleibenden Unfähigkeit, die Muttersprache wieder zu erlernen.
 - III. ist bei zwanzigjährigen Rechtshändern die Fähigkeit, eine Sprache zu erlernen, in der Regel verloren gegangen.
- (A) Nur Ausfall I ist zu erwarten.
(B) Nur Ausfall II ist zu erwarten.
(C) Nur Ausfall III ist zu erwarten.
(D) Nur die Ausfälle I und III sind zu erwarten.
(E) Nur die Ausfälle II und III sind zu erwarten.

Abbildung 6: Itembeispiel Medizinisch-naturwissenschaftliches Grundverständnis (Hänsgen & Spicher, 2010, S. 97f.)

Diagramme und Tabellen

Mit diesem Untertest wird die Analyse- und Interpretationsfähigkeit von Information, die anhand von Diagrammen und Tabellen vorgegeben wird, erhoben (Hänsgen & Spicher, 2010).

Die UntersuchungsteilnehmerInnen in der Studie erhielten zu dem Untertest *Diagramme und Tabellen* insgesamt 16 Items mit einer Bearbeitungszeit von 10 Minuten. Nachfolgend wird ein Itembeispiel dargestellt (siehe Abbildung 7).

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) Milch dem Organismus ihres Konsumenten liefern können.

Milchart	Eiweiss	Fett	Milchzucker	Salze	Energiegehalt
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 Milchzucker wie Buttermilch.
- (B) Vollmilch enthält im Vergleich zur menschlichen Muttermilch etwa die dreifache Menge an Salzen und Eiweiss.
- (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) Der Eiweissgehalt der Milch ist für den Energiegehalt von entscheidender Bedeutung.

Abbildung 7: Itembeispiel Diagramme und Tabellen (Hänsgen & Spicher, 2010, S. 100)

Figuren lernen

Mittels dieses Gedächtnistests wird die Einpräge- und Merkfähigkeit von bestimmten Einzelheiten einer Figur erhoben. In der Einprägephase werden Figuren mit geschwärzten Flächen vorgegeben und nach einer Stunde müssen die UntersuchungsteilnehmerInnen in der Reproduktionsphase den geschwärzten Teil der Figur anführen (Hänsgen & Spicher, 2010).

Ein Itembeispiel ist anhand von Abbildung 8 ersichtlich. In der Studie wurden zu diesem Untertest 8 Items vorgegeben, wobei die Einpräge- und Reproduktionszeit jeweils zwei Minuten betrug und dazwischen ein Zeitraum von 45 Minuten lag.

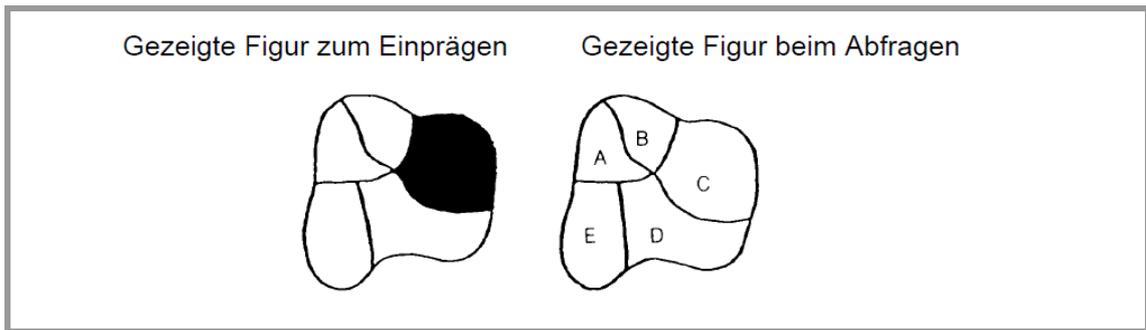


Abbildung 8: Itembeispiel Figuren lernen (Hänsgen & Spicher, 2010, S. 99)

Fakten lernen

Der Untertest *Fakten lernen* stellt entsprechend dem Subtest *Figuren lernen* einen Gedächtnistest dar, der die Einpräge- und Merkfähigkeit von bestimmten Fakten zu Patienten (z.B.: Name, Diagnose) erfasst (Hänsgen & Spicher, 2010).

Die Abbildung 9 zeigt ein Itembeispiel zu diesem Untertest.

Im Rahmen der Untersuchung wurden den SchülerInnen zu diesem Untertest 8 Items vorgegeben, wobei die Einpräge- und Reproduktionszeit jeweils drei Minuten betrug und ein Zeitraum von 44 Minuten zwischen dem Einprägen und Reproduzieren lag.

Lemke, 30 Jahre, Dachdecker, ledig, Schädelbasisbruch

Eine Frage zum obigen Beispiel könnte z.B. lauten:
Der Patient mit dem Schädelbasisbruch ist von Beruf...

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

Abbildung 9: Itembeispiel Fakten lernen (Hänsgen & Spicher, 2010, S. 99)

Abschließend bietet Tabelle 5 einen Überblick zur internen Konsistenz (Cronbach's Alpha) zu den dargestellten EMS-Untertests anhand der Ergebnisse des EMS-Tests 2010 in der Schweiz (Hänsgen & Spicher, 2010) und zu den Schlauchfiguren (Stumpf & Fay, 1983).

Tabelle 5: interne Konsistenz (Cronbach´s Alpha) der Untertests

Untertest	Cronbach´s Alpha	
Quantitative und formale Probleme	.68	(Hänsgen & Spicher, 2010)
Schlauchfiguren	.78	(Stumpf & Fay, 1983)
Medizinisch-naturw. Grundverständnis	.69	(Hänsgen & Spicher, 2010)
Diagramme und Tabellen	.70	(Hänsgen & Spicher, 2010)
Figuren lernen	.69	(Hänsgen & Spicher, 2010)
Fakten lernen	.76	(Hänsgen & Spicher, 2010)

7.2.4 Theory of Mind - Aufgaben

„Unter dem Begriff »Theory of Mind« (ToM) versteht man die Fähigkeit, sich in andere Menschen hineinzusetzen und ihre Gedanken, Motive und Wünsche zu erschließen. ToM setzt die Identifikation und korrekte Bewertung emotionaler Hinweisreize voraus“ (Moritz, Veckenstedt, Randjbar & Vitzthum, 2011, S. 41).

Zur Erhebung der Theory of Mind hinsichtlich einer false belief erster, zweiter und dritter Ordnung wurden drei konstruierte Geschichten (Willinger, Schmöger, Müller & Auff, in Vorbereitung) vorgegeben.

Zu den drei Geschichten müssen die Testpersonen jeweils 13 Fragen beantworten. Das Antwortformat der Items ist zwei- bis vierkategorial und immer nur eine der Antwortmöglichkeiten ist richtig.

7.2.5 Leistungsmotivationsinventar

Das *Leistungsmotivationsinventar* von Schuler und Prochaska (2001) wurde zur Erfassung der berufsbezogenen Leistungsmotivation verwendet. Das Verfahren stellt ein Selbstbeurteilungsverfahren dar, das insgesamt 17 Dimensionen mittels jeweils 10

Items mit siebenstufigem Antwortformat („trifft gar nicht zu“ – „trifft vollständig zu“) erhebt (Schuler & Prochaska, 2001).

Anhand der Abbildung 10 ist ein Itembeispiel des Verfahrens ersichtlich.

	trifft gar nicht zu	trifft vollständig zu
Wenn ich sehe, dass andere mehr können als ich, so ist das ein Ansporn, mich künftig mehr anzustrengen.	①-②-③-④-⑤-⑥-⑦	

Abbildung 10: Itembeispiel Leistungsmotivationsinventar (Schuler & Prochaska, 2001)

In der Studie kamen von den insgesamt 17 Dimensionen zwei Skalen zur Anwendung: Wettbewerbsorientierung und Internalität (Schuler & Prochaska, 2001).

7.2.6 Skala zur Allgemeinen Selbstwirksamkeitserwartung

Die Erhebung der Selbstwirksamkeitserwartung erfolgte mittels der *Skala zur Allgemeinen Selbstwirksamkeitserwartung* von Jerusalem und Schwarzer (1999).

Die Skala wurde auf Basis des Konzepts der Selbstwirksamkeitserwartung von Bandura konstruiert. Ursprünglich bestand diese Skala aus insgesamt 20 Items, jedoch seit 1981 ist das Verfahren im Rahmen der auf 10 Items gekürzten Fassung einsetzbar. Das Erhebungsinstrument wurde über die Jahre hinweg kaum verändert und die Eindimensionalität der Skala konnte anhand mehrerer Untersuchungen (z.B.: Hinz et al., 2006) belegt werden (Schwarzer, 2002).

„Die Skala misst die subjektive Überzeugung, kritische Anforderungssituationen aus eigener Kraft erfolgreich bewältigen zu können“ (Schwarzer, 2002, S. 362).

Das Instrument stellt ein Selbstbeurteilungsverfahren dar, das insgesamt 10 Items mit vierstufigem Antwortformat („stimmt nicht“, „stimmt kaum“, „stimmt eher“, „stimmt genau“) umfasst und eine durchschnittliche Bearbeitungsdauer von vier Minuten in Anspruch nimmt (Schwarzer, 2002). Abbildung 11 stellt ein Itembeispiel der Skala dar.

	Diese Aussage...			
	stimmt nicht	stimmt kaum	stimmt eher	stimmt genau
Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen.	1	2	3	4

Abbildung 11: Itembeispiel Skala zur Allgemeinen Selbstwirksamkeitserwartung (Jerusalem & Schwarzer, 1999)

Die psychometrische Qualität der Skala kann anhand vieler Studien als überaus zufrieden stellend beurteilt werden (Schwarzer, 2002). Die interne Konsistenz (Cronbachs Alpha) der Skala lag in einer Studie zwischen .78 und .79 (Jerusalem & Schwarzer, 1999). Die kriterienbezogene Validität kann belegt werden, beispielsweise können positive Zusammenhänge zur Arbeitszufriedenheit sowie zum dispositionalen Optimismus und negative Zusammenhänge zur Depressivität und zur Ängstlichkeit angenommen werden (Schwarzer, 2002).

Unter Berücksichtigung der Verrechnungsvorschriften (vgl. Schwarzer, 2009) wird im Rahmen der Diplomarbeit die Bildung des Meanscores nur bei jenen Testpersonen durchgeführt, die zumindest sieben der 10 Items beantwortet haben.

7.2.7 Game of Dice Task

Das computerbasierte Erhebungsinstrument *Game of Dice Task* (Brand et al., 2005) wurde zur Erfassung der Risikoverhaltens-Tendenz und zur Überprüfung der Wirksamkeit der Risikoverhalten-Bedingungen eingesetzt.

Dieses Verfahren zeigt auf dem Display einen Würfel und einen Würfelbecher und die UntersuchungsteilnehmerInnen wurden aufgefordert ihr fiktives Startkapital von 1000 Euro innerhalb von 18 Würfeldurchgängen zu erhöhen. Vor jedem Würfeldurchgang können sich die Testpersonen entweder für eine einzelne Zahl oder eine Zahlenkombination (zwei, drei oder vier Zahlen) entscheiden, wobei jede Entscheidung mit einem bestimmten fiktiven Gewinn oder Verlust verbunden ist. Wenn die Testperson auf die gewürfelte Zahl getippt hat, wird der jeweilige Gewinn auf das fiktive Kapital hinzugefügt, bzw. bei einem falschen Tipp abgezogen (Brand et al., 2005).

Anhand der Abbildung 12 wird das Display des Game of Dice Task dargestellt.

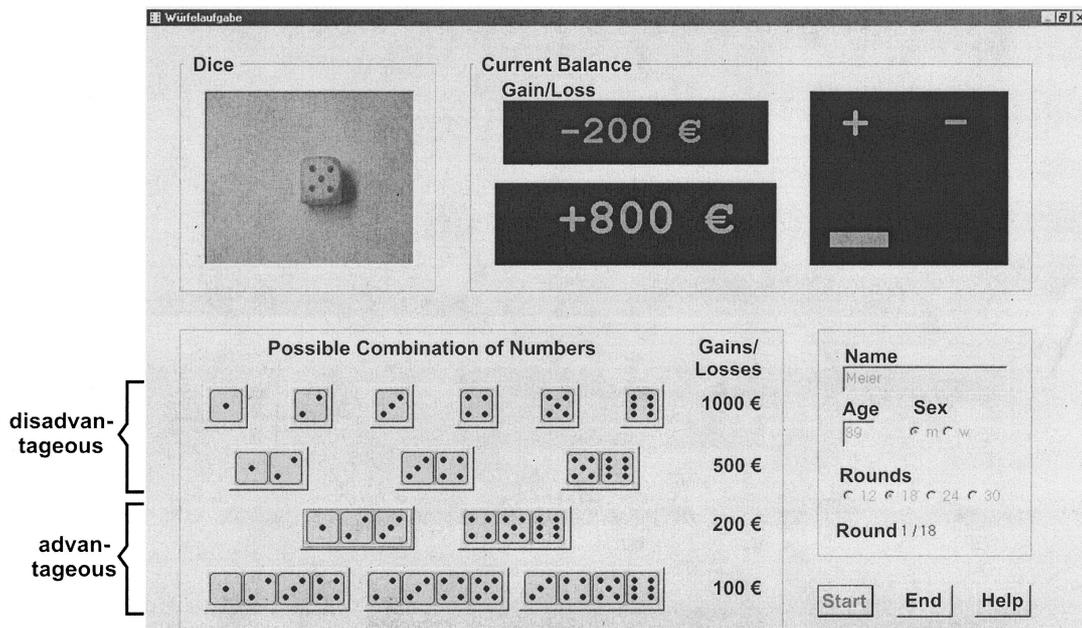


Abbildung 12: Game of Dice Task (Brand et al., 2005, S. 270)

Die Abbildung 12 zeigt, dass die Entscheidung für eine einzelne Zahl an den Gewinn bzw. Verlust von 1000 Euro gebunden ist und die Wahl einer Zahlenkombination führt zu folgenden Gewinn respektive Verlust: zwei Zahlen: 500 Euro, drei Zahlen: 200 Euro und vier Zahlen 100 Euro. Demnach werden Entscheidungen für eine einzelne Zahl oder eine Zahlenkombination aus zwei Zahlen als risikoreich beurteilt, da die Gewinnwahrscheinlichkeit geringer als 50 Prozent ist und die Entscheidung mit einem hohen Gewinn oder mit einem hohen Verlust verbunden ist. Die Entscheidung für eine Zahlenkombination aus drei oder vier Zahlen wird als nicht risikoreich bewertet, da die Gewinnwahrscheinlichkeit bei 50 Prozent und höher liegt sowie geringe Gewinne bzw. Verluste in Zusammenhang stehen (Brand et al., 2005).

8. Untersuchung

In diesem Kapitel werden zunächst die Untersuchungsdurchführung und die statistischen Auswertungsverfahren dargestellt. Daran anschließend erfolgt die Beschreibung der Stichprobe hinsichtlich interessierender sozialstatistischer Merkmale.

8.1 Untersuchungsdurchführung

Nach dem Abschluss der planerischen Vorarbeit und der Fertigstellung des Testheftes wurde das Studienkonzept bei der Ethikkommission der Medizinischen Universität Wien eingereicht. Die Ethikkommission genehmigte die Untersuchung, woraufhin mit den Vorbereitungen einer Vorstudie zur Überprüfung der festgesetzten Bearbeitungszeit (100 Minuten) des Testheftes und der Wirksamkeit der Stereotype Threat-Bedingungen und der Risikoverhalten-Bedingungen, begonnen wurde. An der Vorstudie nahmen insgesamt 31 MedizinstudentInnen (22 Frauen und 9 Männer) im Alter zwischen 20 - 45 Jahren ($M = 23.12$, $SD = 4.62$, $N = 26$ gültige Fälle) teil. Die Untersuchungsdurchführung zeigte, dass die geplante Testzeit von 100 Minuten um etwa +10 Minuten nicht eingehalten werden konnte. Aus diesem Grund konnten nicht, wie geplant, sieben EMS-ähnliche Untertests in das Testheft aufgenommen werden. Die Testbatterie war um den Untertest *Textverständnis* zu verringern. Die Ergebnisse der Voruntersuchung wiesen darauf hin, dass die Wirksamkeit der Stereotype Threat-Bedingungen und der Risikoverhalten-Bedingungen angenommen werden kann. Nach dem Erhalt der Bewilligungen für die wissenschaftliche Untersuchung an den AHS von den Schulräten in Wien, Niederösterreich und dem Burgenland wurde mit den Schulen, die sich bereit erklärten an der Testung teilzunehmen, Erhebungstermine vereinbart. Die erste Testung konnte Ende März 2010 durchgeführt werden. Vor dem jeweils festgelegten Testtermin erhielten die Schulen die Einverständniserklärungen für die Teilnahme an der Studie für die SchülerInnen bzw. die Eltern, wenn die SchülerInnen noch nicht volljährig waren sowie Elternbriefe. Die Durchführung der Untersuchung fand grundsätzlich in den Klassenzimmern der Schulen statt, während eine Schule auch den Turnsaal zur Verfügung stellte und somit in einem Durchgang eine größere Anzahl von SchülerInnen getestet werden konnte. Die Erhebungen nahmen jeweils zwei Unterrichtsstunden in Anspruch und wurden von meinen Kolleginnen und mir durchgeführt. Die Vorgabe der Untersuchungsbedingungen (siehe Untersuchungsplan

und intendierte Stichprobe) erfolgte gemäß der Planung, mit der Ausnahme, dass das *Verstehen von komplexen Texten* aus den Instruktionen zu den Stereotype Threat-Bedingungen Aktivierung und Negativ-Aktivierung entfiel, da ja der Untertest *Textverständnis* aufgrund der Voruntersuchung aus der Testbatterie gestrichen wurde. Gemäß dem Untersuchungsplan sollte die mögliche Mediatorvariable Ängstlichkeit neben der Selbstwirksamkeit und der Leistungsmotivation nach der Bearbeitung der Theory of Mind – Aufgaben erhoben werden, jedoch wegen der zeitlichen Begrenzung der Testung, konnte diese Variable in die Studie nicht einbezogen werden. Die Hormonmessungen wurden vor und nach der Bearbeitung des EMS-ähnlichen Test durchgeführt, konnten aber hinsichtlich der zur Verfügung stehenden finanziellen Mittel nicht bei allen Testpersonen erhoben werden. Nach der Bearbeitung des Testheftes wurde den UntersuchungsteilnehmerInnen das Erhebungsinstrument *Game of Dice Task* (Brand et al., 2005) am Laptop vorgegeben. Aufgrund der beschränkten Testzeit von zwei Unterrichtsstunden konnte diese Erhebung nicht bei allen Testpersonen erfolgen. Die Testungen konnten mit Ende Jänner 2011 abgeschlossen werden, woraufhin mit der Dateneingabe und Auswertung begonnen wurde.

8.2 Auswertungsverfahren

Die deskriptive und inferenzstatistische Auswertung der durchgeführten Untersuchung erfolgt mittels IBM SPSS Statistics 19. Im Folgenden werden die statistischen Auswertungsverfahren dargestellt:

Kontingenztafel und χ^2 -Test

Anhand der Kontingenztafel (Kreuztabelle) können Zusammenhänge zwischen nominalskalierten Variablen veranschaulicht und untersucht werden. Wird die Vermutung eines Zusammenhangs durch die Kontingenztafel gestützt, kann mittels Chi-Quadrat-Test (χ^2 -Test) analysiert werden, ob eine zufällige Assoziation oder ein systematischer Zusammenhang vorliegt (Backhaus, Erichson, Plinke & Weiber, 2011).

Produkt-Moment-Korrelation nach Pearson

Die Produkt-Moment-Korrelation nach Pearson prüft, ob zwischen zwei Variablen ein linearer Zusammenhang existiert. Für die Verwendung dieses Verfahrens sind intervallskalierte Daten sowie Normalverteilung der Daten erforderlich. Mittels des

Korrelationskoeffizient nach Pearson (r) wird die Richtung und die Enge des Zusammenhangs zwischen zwei Variablen aufgezeigt: Werte von 0.1 bis 0.29 weisen auf einen geringen, Werte von 0.30 bis 0.49 auf einen mittleren und Werte ≥ 0.50 auf einen hohen Zusammenhang hin (Bortz & Döring, 2005).

Univariate Varianzanalyse (ANOVA)

Mit diesem statistischen Auswertungsverfahren wird die Wirkung von einer oder mehreren nominalskalierten, unabhängigen Variablen (UV) auf eine abhängige Variable (AV) mit zumindest Intervallskalenniveau, analysiert. Die unabhängigen Variablen nennt man Faktoren und die jeweiligen Ausprägungen Faktorstufen. Das Verfahren wird je nach Anzahl der unabhängigen Variablen als einfaktorielle (1 UV) Varianzanalyse, zweifaktorielle (2 UV) Varianzanalyse, usw. bezeichnet (Backhaus et al., 2011).

Für die Anwendung dieses Verfahrens müssen folgende Voraussetzungen gegeben sein: Intervallskalenniveau der abhängigen Variable, Normalverteilung der AV innerhalb jeder Gruppe und Homogenität der Varianzen (Field, 2009).

Multivariate Varianzanalyse (MANOVA)

Mittels MANOVA wird der Einfluss von einer oder mehreren nominalskalierten, unabhängigen Variablen, auf mehrere abhängige Variablen (intervallskaliert), gleichzeitig überprüft. Die Anwendung der multivariaten Varianzanalyse ist an die Erfüllung der nachfolgenden Voraussetzungen gebunden: Intervallskalenniveau der abhängigen Variablen, multivariate Normalverteilung der abhängigen Variablen innerhalb der Gruppen, Homogenität der Varianzen und Homogenität der Varianz-Kovarianz Matrizen der abhängigen Variablen (Field, 2009).

Levene´s-Test

Mittels Levene´s-Test wird die Nullhypothese, dass die Varianzen in verschiedenen Gruppen gleich sind („Homogenität der Varianzen“), überprüft. Ein signifikantes Ergebnis bedeutet, dass sich die Varianzen signifikant unterscheiden und die Voraussetzung der Homogenität der Varianzen nicht erfüllt ist, wobei bei einem nicht signifikanten Ergebnis die Homogenität der Varianzen gegeben ist (Field, 2009).

Box's M Test

Mittels Box's M Test wird die Homogenität der Varianz-Kovarianz-Matrizen der abhängigen Variablen zwischen den Gruppen verglichen. Ein signifikantes Ergebnis weist auf die Ungleichheit der Varianz-Kovarianz-Matrizen hin, während bei einem nicht signifikanten Ergebnis die Homogenität der Varianz-Kovarianz Matrizen der abhängigen Variablen zwischen den Gruppen angenommen werden kann (Field, 2009).

A-priori Kontraste

Ein signifikantes Ergebnis, beispielsweise einer ANOVA, erfordert weitere Analysen um zu prüfen welche Gruppen sich unterscheiden. Wenn bereits vor der Analyse spezifische Hypothesen bezüglich vermuteter Mittelwertsunterschiede der Gruppen formuliert wurden, erfolgt mittels a-priori Kontrast der Vergleich jeweils von zwei Gruppen auf signifikante Unterschiede (Field, 2009).

Multivariate Kovarianzanalyse (MANCOVA)

Erweiternd zur MANOVA bezieht die multivariate Kovarianzanalyse eine oder mehrere metrische, personenbezogene Kovariate(n) in die Berechnungen mit ein (Backhaus et al., 2011; Bortz & Döring, 2005). Eine Kovariate stellt eine potentielle Störvariable dar, deren vermutlicher Einfluss auf die abhängige(n) Variable(n) mittels der (multivariaten) Kovarianzanalyse herauspartialisiert wird (Bortz & Döring, 2005). Zusätzlich zu den Voraussetzungen der MANOVA muss bei der MANCOVA auch die Homogenität der Regressionsgeraden gegeben sein (Field, 2009).

t-Test für unabhängige Stichproben

Mittels t-Test für unabhängige Stichproben werden die Mittelwerte von zwei unabhängigen Stichproben aus verschiedenen Populationen miteinander verglichen. Es wird überprüft, ob sich die Mittelwerte der zwei Gruppen signifikant voneinander unterscheiden. Die Anwendung dieses Verfahrens ist an folgende Voraussetzungen gebunden: Normalverteilung, intervallskalierte Daten, Homogenität der Varianzen und Unabhängigkeit der Testwerte (Field, 2009).

8.3 Stichprobenbeschreibung

Nachfolgend wird die Stichprobe hinsichtlich interessierender sozialstatistischer Merkmale beschrieben und dargestellt.

An den Testungen, die im Zeitraum von Ende März 2010 bis Ende Jänner 2011 durchgeführt wurden nahmen insgesamt 1093 SchülerInnen der 12. Schulstufe von Gymnasien unterschiedlicher Standorte und Ausbildungszweige teil.

8.3.1 Geschlecht

Von den 1093 UntersuchungsteilnehmerInnen waren 684 (62,6 %) weiblich und 409 (37,4 %) männlich. Die Abbildung 13 verdeutlicht die Geschlechterverteilung in der Stichprobe.

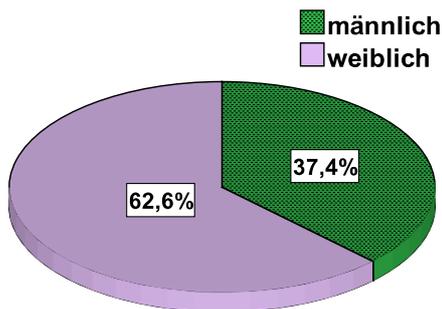


Abbildung 13: Geschlecht (N = 1093 gültige Fälle)

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 69.190$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 6). Es kann ein Verteilungsunterschied im Geschlecht beobachtet werden.

Tabelle 6: Chi-Quadrat-Test Geschlecht

	Geschlecht
Chi-Quadrat	69.190 ^a
df	1
Asymptotische Signifikanz	$\leq .0001$

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

8.3.2 Alter

Die Altersspanne der TestteilnehmerInnen (N = 1086 gültige Fälle) liegt bei 16–21 Jahren. Das Durchschnittsalter der SchülerInnen betrug beim Untersuchungszeitpunkt 17.57 Jahre ($SD = 0.76$). Von den insgesamt 1093 UntersuchungsteilnehmerInnen gaben sieben SchülerInnen das Alter nicht an. Tabelle 7 stellt die Häufigkeitsverteilung des Alters der TestteilnehmerInnen dar.

Tabelle 7: Häufigkeitsverteilung Alter (N = 1086 gültige Fälle)

Alter				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	16	13	1,2	1,2
	17	580	53,1	53,4
	18	388	35,5	35,7
	19	80	7,3	7,4
	20	21	1,9	1,9
	21	4	,4	,4
	Gesamt	1086	99,4	100,0
Fehlend	System	7	,6	
Gesamt		1093	100,0	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(5) = 1643.116$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 8). Es kann ein Verteilungsunterschied im Alter angenommen werden.

Tabelle 8: Chi-Quadrat-Test Alter

	Alter
Chi-Quadrat	1643.116 ^a
df	5
Asymptotische Signifikanz	$\leq .0001$

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

8.3.3 Schulstandort und Staatsbürgerschaft

Die Untersuchung fand bei insgesamt 24 allgemeinbildenden höheren Schulen in drei verschiedenen Bundesländern statt. Der Großteil der 1093 getesteten SchülerInnen besuchte Gymnasien in Wien (72,7 %) und eine geringere Anzahl Gymnasien in Niederösterreich (20,0 %) oder im Burgenland (7,3 %).

Anhand der Abbildung 14 ist ersichtlich, dass von 1050 gültigen Fällen, die überwiegende Mehrheit der UntersuchungsteilnehmerInnen (94,5 %) eine österreichische Staatsbürgerschaft besitzt. Eine ausländische Staatsbürgerschaft (u.a. Serbien, Ungarn, Kroatien, Polen, Slowakei, Japan, Indien, Philippinen) haben 5,5 % der SchülerInnen. Von den 1093 UntersuchungsteilnehmerInnen führten 43 SchülerInnen die Staatsbürgerschaft nicht an.

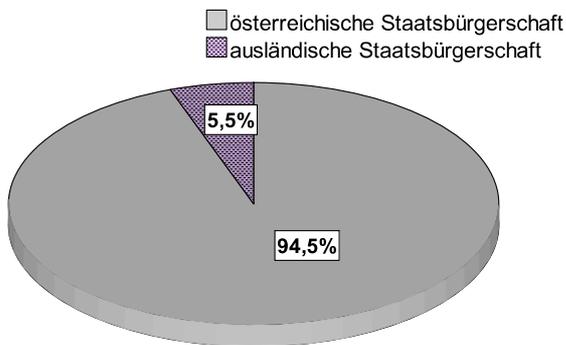


Abbildung 14: Staatsbürgerschaft (N = 1050 gültige Fälle)

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(5) = 4384.784$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 9). Es kann ein Verteilungsunterschied in der Staatsbürgerschaft beobachtet werden.

Tabelle 9: Chi-Quadrat-Test Staatsbürgerschaft

	Staatsbürgerschaft
Chi-Quadrat	4384.784 ^a
df	5
Asymptotische Signifikanz	$\leq .0001$

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 180,5.

8.3.4 Ausbildungszweig

Mittels Tabelle 10 wird die Häufigkeitsverteilung der Ausbildungszweige, die von den UntersuchungsteilnehmerInnen absolviert werden dargelegt. 13 SchülerInnen gaben den Ausbildungszweig nicht an.

Tabelle 10: Häufigkeitsverteilung Ausbildungszweig (N = 1080 gültige Fälle)

Ausbildungszweig				
		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	,1	,1
	Wirtschaftlich	27	2,5	2,5
	Humanistisch	20	1,8	1,9
	Informatik	20	1,8	1,9
	Sonstiger Ausbildungszweig	68	6,2	6,3
	Sport	97	8,9	9,0
Gesamt		1080	98,8	100,0
Fehlend	System	13	1,2	
Gesamt		1093	100,0	

Von den 1080 gültigen Fällen absolviert die Mehrzahl der SchülerInnen (26,7%) einen sprachlichen Ausbildungszweig, gefolgt von dem naturwissenschaftlichen (22,9 %) und dem musikalischen Ausbildungszweig (18,4 %). Die Abbildung 15 veranschaulicht die Verteilung der Ausbildungszweige anhand eines Kreisdiagramms.

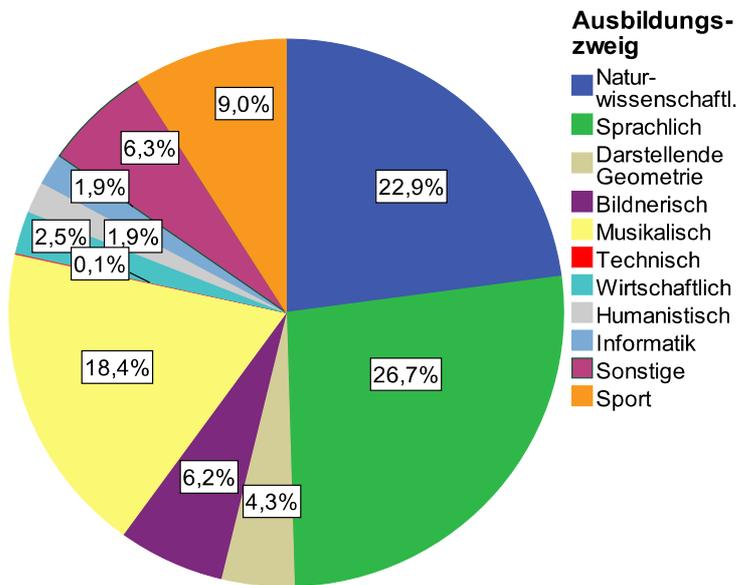


Abbildung 15: Ausbildungszweig (N = 1080 gültige Fälle)

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(11) = 1184.621$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 11). Es kann ein Verteilungsunterschied im Ausbildungszweig angenommen werden.

Tabelle 11: Chi-Quadrat-Test Ausbildungszweig

	Ausbildungszweig
Chi-Quadrat	1184.621 ^a
df	11
Asymptotische Signifikanz	$\leq .0001$

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

8.3.5 Wiederholung einer Schulstufe

Die Häufigkeitsverteilung der Wiederholung von zumindest einer Schulstufe wird anhand von Tabelle 12 dargestellt. Es ist ersichtlich, dass von 1087 gültigen Fällen, 166 Testpersonen (15,3 %) zumindest eine Schulstufe wiederholt haben.

Tabelle 12: Häufigkeitsverteilung Wiederholung einer Schulstufe (N = 1087 gültige Fälle)

Wiederholung einer Schulstufe				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	nein	921	84,3	84,7
	ja	166	15,2	15,3
	Gesamt	1087	99,5	100,0
Fehlend	System	6	,5	
Gesamt		1093	100,0	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 524.402$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 13). Es kann ein Verteilungsunterschied im Wiederholen von zumindest einer Schulstufe angenommen werden.

Tabelle 13: Chi-Quadrat-Test Wiederholung einer Schulstufe

	Wiederholung einer Schulstufe
Chi-Quadrat	524.402 ^a
df	1
Asymptotische Signifikanz	$\leq .0001$

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

Die Kontingenztafel (Tabelle 14) stellt das Wiederholen zumindest einer Schulstufe in Abhängigkeit vom Geschlecht dar. Anhand der Kontingenztafel ist erkennbar, dass 19,1 % der männlichen Testteilnehmer und 13 % der weiblichen Testpersonen zumindest eine Schulstufe wiederholt haben.

Tabelle 14: Kontingenztabelle Geschlecht und Wiederholung zumindest einer Schulstufe

Geschlecht * Wiederholung einer Schulstufe Kreuztabelle					
			Wiederholung einer Schulstufe		Gesamt
			nein	ja	
Geschlecht	männlich	Anzahl	326	77	403
		Erwartete Anz.	341.5	61.5	403.0
		% innerhalb	80,9%	19,1%	100,0%
		Stand. Residuen	-.8	2.0	
	weiblich	Anzahl	595	89	684
		Erwartete Anz.	579.5	104.5	684.0
		% innerhalb	87,0%	13,0%	100,0%
		Stand. Residuen	.6	-1.5	
Gesamt		Anzahl	921	166	1087
		Erwartete Anz.	921.0	166.0	1087.0
		% innerhalb	84,7%	15,3%	100,0%

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 7.281$, $p = .007$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied im Wiederholen einer Schulstufe in der Oberstufe in Abhängigkeit vom Geschlecht beobachtet werden. Mittels Tabelle 14 wird deutlich, dass männliche Testteilnehmer häufiger eine Schulstufe wiederholen als die weiblichen Testpersonen.

8.3.6 Ausbildung der Eltern

Anhand Tabelle 15 wird die Häufigkeitsverteilung der höchsten abgeschlossenen Ausbildung der Mütter der Testpersonen (N = 1063 gültige Fälle) dargelegt. Die beiden stärksten Anteile der höchsten absolvierten Ausbildung der Mütter der UntersuchungsteilnehmerInnen stellen mit 31,3 % die Matura und mit 31,1 % die Universität dar. Der Abschluss einer berufsbildenden mittleren Schule oder einer Lehre nimmt mit 26,0 % den drittstärksten Anteil der höchsten absolvierten Ausbildung der Mütter ein. 30 SchülerInnen führten die höchste abgeschlossene Ausbildung ihrer Mütter nicht an.

Tabelle 15: Häufigkeitsverteilung höchste abgeschlossenen Ausbildung der Mütter (N = 1063 gültige Fälle)

Ausbildung der Mutter				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	Ohne Schulabschluss	13	1,2	1,2
	Pflichtschule	77	7,0	7,2
	Berufsbildende mittlere Schule, Lehre	276	25,3	26,0
	Matura	333	30,5	31,3
	Fachhochschule	27	2,5	2,5
	Universität	331	30,3	31,1
	Sonstiges	6	,5	,6
	Gesamt	1063	97,3	100,0
Fehlend	System	30	2,7	
Gesamt		1093	100,0	

Die Abbildung 16 verdeutlicht die Verteilung der höchsten absolvierten Ausbildung der Mütter der UntersuchungsteilnehmerInnen anhand eines Kreisdiagramms.

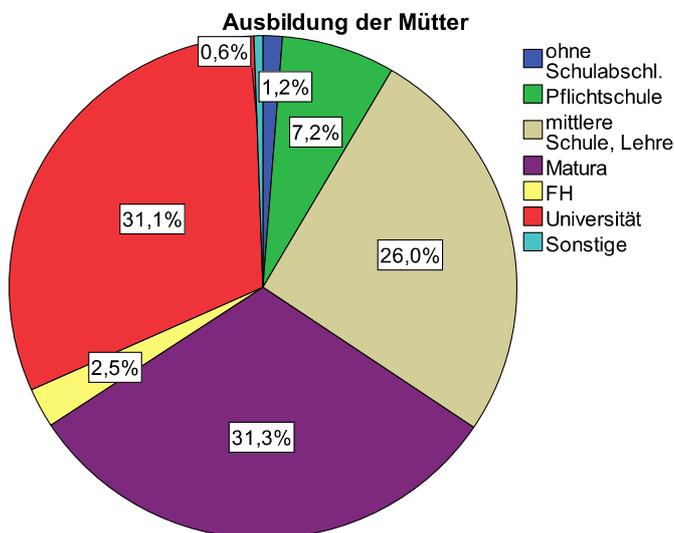


Abbildung 16: Höchste abgeschlossene Ausbildung der Mütter (N = 1063 gültige Fälle)

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(7) = 1134.916$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann

ein Verteilungsunterschied in der höchsten abgeschlossenen Ausbildung der Mütter der Testpersonen angenommen werden (siehe Tabelle 16).

Tabelle 16: Chi-Quadrat-Test höchste abgeschlossene Ausbildung der Mütter

	Ausbildung der Mütter
Chi-Quadrat	1134.916 ^a
df	7
Asymptotische Signifikanz	≤ .0001

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 136,6.

Die Häufigkeitsverteilung der höchsten abgeschlossenen Ausbildung der Väter der UntersuchungsteilnehmerInnen wird mittels Tabelle 17 dargestellt. Von den 1093 SchülerInnen gaben 45 Testpersonen die Ausbildung ihrer Väter nicht an.

Tabelle 17: Häufigkeitsverteilung höchste abgeschlossene Ausbildung der Väter (N = 1048 gültige Fälle)

Ausbildung des Vaters				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	Ohne Schulabschluss	9	,8	,9
	Pflichtschule	83	7,6	7,9
	Berufsbildende mittlere Schule, Lehre	255	23,3	24,3
	Matura	277	25,3	26,4
	Fachhochschule	20	1,8	1,9
	Universität	402	36,8	38,4
	Sonstiges	2	,2	,2
	Gesamt	1048	95,9	100,0
Fehlend	System	45	4,1	
Gesamt		1093	100,0	

Die Mehrzahl der Väter der SchülerInnen mit 38,4 % (N = 1048 gültige Fälle) absolvierte eine Universität. Die Matura mit 26,4% und der Abschluss einer berufsbildenden mittleren Schule oder einer Lehre mit 24,3 % stellen die zweit- bzw. drittstärksten Anteile der höchsten absolvierten Ausbildung der Väter dar. Abbildung 17

veranschaulicht die Verteilung der höchsten abgeschlossenen Ausbildung der Väter der Testpersonen.

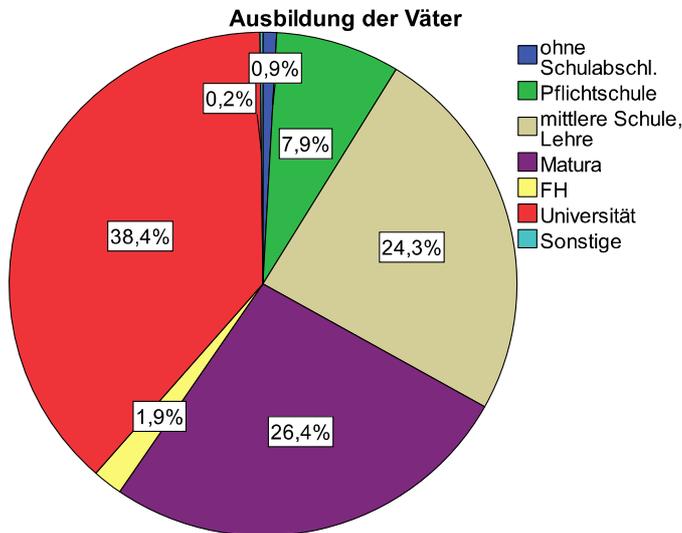


Abbildung 17: Höchste abgeschlossene Ausbildung der Väter (N = 1048 gültige Fälle)

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(7) = 1196.164$, $p \leq .0001$ ein signifikantes Ergebnis (siehe Tabelle 18). Es kann ein Verteilungsunterschied in der höchsten abgeschlossenen Ausbildung der Väter der UntersuchungsteilnehmerInnen beobachtet werden.

Tabelle 18: Chi-Quadrat-Test höchste abgeschlossene Ausbildung der Väter

	Ausbildung der Väter
Chi-Quadrat	1196.164 ^a
df	7
Asymptotische Signifikanz	$\leq .0001$

a. Bei 0 Zellen (0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 136,6.

8.3.7 Intention, Medizin zu studieren

Mittels Tabelle 19 wird die Häufigkeitsverteilung der Intention, ein Medizinstudium zu beginnen, dargestellt. Die Tabelle zeigt, dass 187 Testpersonen (17,5 %) Medizin studieren möchten (N= 1068 gültige Fälle).

Tabelle 19: Häufigkeitsverteilung Intention, Medizin zu studieren

Intention, Medizin zu studieren				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	nein	881	80,6	82,5
	ja	187	17,1	17,5
	Gesamt	1068	97,7	100,0
Fehlend	System	25	2,3	
Gesamt		1093	100,0	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 450.970$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Intention, ein Medizinstudium zu beginnen beobachtet werden (siehe Tabelle 20).

Tabelle 20: Chi-Quadrat-Test Intention, Medizin zu studieren

	Intention, Medizin zu studieren
Chi-Quadrat	450.970 ^a
df	1
Asymptotische Signifikanz	$\leq .0001$

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

Auf Grundlage der Kontingenztafel (Tabelle 21) wird die Intention, nach der Matura ein Medizinstudium zu beginnen in Abhängigkeit vom Geschlecht dargestellt. Die Tabelle zeigt, dass von den 187 Testpersonen, die angaben nach der Matura die Intention für den Beginn eines Medizinstudiums zu haben, 125 Schülerinnen (66,8 %) und 62 Schüler (33,2 %) sind.

Tabelle 21: Kontingenztafel Geschlecht und die Intention, Medizin zu studieren

Geschlecht *Intention, Medizin zu studieren Kreuztabelle					
			Medizinstudium		Gesamt
			nein	ja	
Geschlecht	männlich	Anzahl	334	62	396
		Erwartete Anz.	326.7	69.3	396.0
		% innerhalb	84,3%	15,7%	100,0%
		Stand. Residuen	.4	-.9	
	weiblich	Anzahl	547	125	672
		Erwartete Anz.	554.3	117.7	672.0
		% innerhalb	81,4%	18,6%	100,0%
		Stand. Residuen	-.3	.7	
Gesamt	Anzahl	881	187	1068	
	Erwartete Anz.	881.0	187.0	1068.0	
	% innerhalb	82,5%	17,5%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 1.496$, $p = .221$ ein nicht signifikantes Ergebnis. Es kann kein Verteilungsunterschied in der Intention, ein Medizinstudium zu beginnen in Abhängigkeit vom Geschlecht beobachtet werden.

8.3.8 Befassung mit dem EMS

Die bisherige Befassung mit dem EMS (z.B.: Testaufgaben des EMS bearbeitet, Medizinstudenten befragt) in Abhängigkeit vom Geschlecht ist anhand Tabelle 22 ersichtlich. Es zeigt sich, dass sich von insgesamt 1085 UntersuchungsteilnehmerInnen 203 SchülerInnen (18,7 %) bereits mit dem EMS auseinandergesetzt haben. Von diesen 203 SchülerInnen teilten 142 weibliche Testpersonen (70,0 %) und 61 männliche Testteilnehmer (30,0 %) eine bereits erfolgte Befassung mit dem EMS mit.

Tabelle 22: Kontingenztafel Geschlecht und die Befassung mit dem EMS

Geschlecht * EMS Befassung Kreuztabelle					
			Befassung mit EMS		Gesamt
			nein	ja	
Geschlecht	männlich	Anzahl	341	61	402
		Erwartete Anz.	326.8	75.2	402.0
		% innerhalb	84,8%	15,2%	100,0%
		Stand. Residuen	.8	-1.6	
	weiblich	Anzahl	541	142	683
		Erwartete Anz.	555.2	127.8	683.0
		% innerhalb	79,2%	20,8%	100,0%
		Stand. Residuen	-.6	1.3	
Gesamt	Anzahl	882	203	1085	
	Erwartete Anz.	882.0	203.0	1085.0	
	% innerhalb	81,3%	18,7%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 5.249$, $p = .022$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Befassung mit dem EMS in Abhängigkeit vom Geschlecht beobachtet werden. Anhand der standardisierten Residuen wird deutlich, dass sich männliche Schüler mit 15,2 % weniger oft als Schülerinnen mit 20,8 % mit dem EMS befassen.

Darüber hinaus wird im Rahmen der Stichprobenbeschreibung dargelegt, ob es auch einen Verteilungsunterschied in der Befassung mit dem EMS in Abhängigkeit von der Intention, ein Medizinstudium zu beginnen, gibt.

Die Tabelle 23 zeigt, dass sich von 187 Testpersonen, die ein Medizinstudium beginnen möchten, 117 SchülerInnen (62,6 %) bereits mit dem EMS auseinandergesetzt haben und 70 Testteilnehmer (37,4 %) noch keine Befassung mit dem EMS äußerten. Von den 879 UntersuchungsteilnehmerInnen, die kein Medizinstudium intendieren gaben 80 Testpersonen (9,1 %) eine erfolgte Befassung und 799 SchülerInnen (90,1 %) keine Auseinandersetzung mit dem EMS an.

Tabelle 23: Kontingenztafel die Intention, Medizin zu studieren und die Befassung mit dem EMS

Intention, Medizin zu studieren * Befassung mit EMS Kreuztabelle					
			Befassung mit EMS		Gesamt
			nein	ja	
Intention, Medizin zu studieren	nein	Anzahl	799	80	879
		Erwartete Anz.	716.6	162.4	879.0
		% innerhalb	90,9%	9,1%	100,0%
		Stand. Residuen	3.1	-6.5	
	ja	Anzahl	70	117	187
		Erwartete Anz.	152.4	34.6	187.0
		% innerhalb	37,4%	62,6%	100,0%
		Stand. Residuen	-6.7	14.0	
Gesamt		Anzahl	869	197	1066
		Erwartete Anz.	869.0	197.0	1066.0
		% innerhalb	81,5%	18,5%	100,0%

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(1) = 292.584$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Befassung mit dem EMS in Abhängigkeit von der Intention, ein Medizinstudium zu beginnen, angenommen werden.

8.3.9 Verteilung der UntersuchungsteilnehmerInnen in den Untersuchungsbedingungen

Im Folgenden wird gezeigt, dass die Zuweisung der SchülerInnen im Rahmen des Untersuchungsdesigns mit dem festen Faktor Geschlecht und den beiden experimentellen Faktoren Stereotype Threat und Risikoverhalten auf die so entstandenen 24 (= 2*4*3) Faktorstufenkombinationen gleichmäßig erfolgt ist.

Anhand von Tabelle 24 ist zunächst die Zuweisung der UntersuchungsteilnehmerInnen zu den Stereotype Threat-Bedingungen ersichtlich.

Tabelle 24: Zuweisung der UntersuchungsteilnehmerInnen zu den Stereotype Threat - Bedingungen

Stereotype Threat - Bedingungen		
	Häufigkeit	Prozent
Aktivierung	293	26,8
Neutrale Bedingung	257	23,5
Aufklärung	280	25,6
Negativ-Aktivierung	263	24,1
Gesamt	1093	100,0

Die Häufigkeitsverteilung der TestteilnehmerInnen auf die vier Stereotype Threat-Bedingungen unterliegt unter Berücksichtigung des Erwartungswerts 273.3 mit $\chi^2(3) = 2.945$, $p = .400$ annähernd einer Gleichverteilung (siehe Tabelle 25).

Tabelle 25: Chi-Quadrat-Test Stereotype Threat-Bedingungen

	Stereotype Threat
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.

Anhand von Tabelle 26 ist die Zuweisung der UntersuchungsteilnehmerInnen zu den Bedingungen im Risikoverhalten ersichtlich.

Tabelle 26: Zuweisung der UntersuchungsteilnehmerInnen zu den Risikoverhalten - Bedingungen

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

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt unter Berücksichtigung des Erwartungswerts 364.3 mit $\chi^2(2) = 0.661$, $p = .719$ ein nicht signifikantes Ergebnis, so dass eine annähernde Gleichverteilung der UntersuchungsteilnehmerInnen auf die drei Risikoverhalten-Bedingungen angenommen werden kann (siehe Tabelle 27).

Tabelle 27: Chi-Quadrat-Test Risikoverhalten-Bedingungen

	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.

Geschlechterverteilung in den Stereotype Threat-Bedingungen

Tabelle 28 zeigt die Verteilung der Testpersonen auf die Stereotype Threat-Bedingungen in Abhängigkeit vom Geschlecht.

Tabelle 28: Kontingenztafel Geschlecht und Stereotype Threat - Bedingungen

Geschlecht * Stereotype Threat Kreuztabelle							
		Stereotype Threat – Bedingungen					Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung		
Geschlecht	männlich	Anzahl	123	85	112	89	409
		Erwartete Anz.	109.6	96.2	104.8	98.4	409.0
		% innerhalb	30,1%	20,8%	27,4%	21,8%	100,0%
		Stand. Residuen	1.3	-1.1	.7	-.9	
	weiblich	Anzahl	170	172	168	174	684
		Erwartete Anz.	183.4	160.8	175.2	164.6	684.0
		% innerhalb	24,9%	25,1%	24,6%	25,4%	100,0%
		Stand. Residuen	-1.0	.9	-.5	.7	
Gesamt	Anzahl	293	257	280	263	1093	
	Erwartete Anz.	293.0	257.0	280.0	263.0	1093.0	
	% innerhalb	26,8%	23,5%	25,6%	24,1%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(3) = 6.909$, $p = .075$ ein nicht signifikantes Ergebnis. Es kann kein Verteilungsunterschied in den Stereotype Threat-Bedingungen in Abhängigkeit vom Geschlecht angenommen werden. Somit kann die Zuweisung der UntersuchungsteilnehmerInnen in die vier Stereotype Threat-Bedingungen als unabhängig vom Geschlecht angesehen werden.

Geschlechterverteilung in den Risikoverhalten-Bedingungen

Anhand von Tabelle 29 ist die Verteilung der TestteilnehmerInnen auf die Risikoverhalten-Bedingungen in Abhängigkeit vom Geschlecht ersichtlich.

Tabelle 29: Kontingenztafel Geschlecht und Risikoverhalten - Bedingungen

Geschlecht * Risikoverhalten Kreuztabelle						
			Risikoverhalten - Bedingungen			Gesamt
			raten	ohne Instruktion	sicher	
Geschlecht	männlich	Anzahl	145	131	133	409
		Erwartete Anz.	139.6	137.7	131.7	409.0
		% innerhalb	35,5%	32,0%	32,5%	100,0%
		Stand. Residuen	.5	-.6	.1	
	weiblich	Anzahl	228	237	219	684
		Erwartete Anz.	233.4	230.3	220.3	684.0
		% innerhalb	33,3%	34,6%	32,0%	100,0%
		Stand. Residuen	-.4	.4	-.1	
Gesamt	Anzahl	373	368	352	1093	
	Erwartete Anz.	373.0	368.0	352.0	1093.0	
	% innerhalb	34,1%	33,7%	32,2%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(2) = 0.878$, $p = .645$ ein nicht signifikantes Ergebnis. Es kann kein Verteilungsunterschied in den Risikoverhalten-Bedingungen in Abhängigkeit vom Geschlecht beobachtet werden. Die Zuteilung der UntersuchungsteilnehmerInnen auf die Risikoverhalten-Bedingungen erfolgte unabhängig vom Geschlecht.

Verteilung der SchülerInnen in den beiden experimentellen Faktoren

Die Verteilung der TestteilnehmerInnen auf die Risikoverhalten-Bedingungen und den Stereotype Threat-Bedingungen wird in Tabelle 30 angeführt.

Tabelle 30: Kontingenztafel Stereotype Threat - Bedingungen und Risikoverhalten - Bedingungen

Stereotype Threat * Risikoverhalten Kreuztabelle						
			Risikoverhalten - Bedingungen			Gesamt
			raten	ohne Instruktion	sicher	
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	93	108	92	293
		Erwartete Anz.	100.0	98.6	94.4	293.0
		% innerhalb	31,7%	36,9%	31,4%	100,0%
		Stand. Residuen	-.7	.9	-.2	
	Neutral	Anzahl	90	87	80	257
		Erwartete Anz.	87.7	86.5	82.8	257.0
		% innerhalb	35,0%	33,9%	31,1%	100,0%
		Stand. Residuen	.2	.1	-.3	
	Aufklärung	Anzahl	104	92	84	280
		Erwartete Anz.	95.6	94.3	90.2	280.0
		% innerhalb	37,1%	32,9%	30,0%	100,0%
		Stand. Residuen	.9	-.2	-.7	
	Negativ-Aktivierung	Anzahl	86	81	96	263
		Erwartete Anz.	89.8	88.5	84.7	263.0
		% innerhalb	32,7%	30,8%	36,5%	100,0%
		Stand. Residuen	-.4	-.8	1.2	
Gesamt	Anzahl	373	368	352	1093	
	Erwartete Anz.	373.0	368.0	352.0	1093.0	
	% innerhalb	34,1%	33,7%	32,2%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 5.121$, $p = .528$ ein nicht signifikantes Ergebnis. Es kann kein Verteilungsunterschied in den Risikoverhalten-Bedingungen in Abhängigkeit der Stereotype Threat-Bedingungen beobachtet werden. Die Zuweisung der SchülerInnen auf die Untersuchungsbedingungen kann als unabhängig von den Stereotype Threat * Risikoverhalten Faktorstufenkombinationen angesehen werden.

9. Ergebnisse

In diesem Kapitel werden zunächst die Deskriptivstatistik und die Koeffizienten zur Reliabilität der EMS-ähnlichen Untertests sowie die Lösungswahrscheinlichkeiten der Aufgaben und die Itemtrennschärfen dargestellt. Weiters erfolgt die Darlegung der Kennwerte und Reliabilitätskoeffizienten zur Selbstwirksamkeitserwartung und zum *Game of Dice Task* (GDT) (Brand et al., 2005). Danach werden die Überprüfung der Wirksamkeit der Stereotype Threat-Untersuchungsbedingungen und die Hypothesenbezogenen Ergebnisse dargestellt.

9.1 Kennwerte der Erhebungsinstrumente

9.1.1 EMS-ähnliche Untertests

Die Testkennwerte zu den EMS-ähnlichen Subtests können Tabelle 31 entnommen werden. Anhand dieser Tabelle werden die durchschnittlich erzielten Punkte der SchülerInnen pro Untertest (M) und die maximal richtig gelösten Items eines Untertests (Max.) aufgezeigt. Beim EMS-ähnlichen Untertest Medizinisch-naturwissenschaftliches Grundverständnis beispielsweise, erzielten die Testpersonen durchschnittlich 2.81 Punkte und von den insgesamt 16 Items in diesem Untertest wurden maximal 9 richtig gelöst. Weiters werden die Werte der Schiefe und der Kurtosis dargelegt. Gemäß Field (2009) kann von einer annähernden Normalverteilung der Daten ausgegangen werden, wenn die Werte der Schiefe und Kurtosis nahe bei Null liegen. Da die Werte nicht wesentlich von Null abweichen, kann angenommen werden, dass die Leistungen der SchülerInnen in den EMS-ähnlichen Untertests jeweils einer Normalverteilung unterliegen.

Tabelle 31: Deskriptivstatistik EMS-ähnliche Subtests (N = 1093)

EMS-ähnliche Untertests	M	SD	Schiefe	Kurtosis	Min.	Max.
Quantitative u. formale Probleme	3.00	1.68	.485	.155	0	10
Schlauchfiguren	10.19	3.45	-.168	-.283	0	20
Medizinisch-naturw. Grundverständnis	2.81	1.69	.482	-.150	0	9
Diagramme u. Tabellen	3.01	1.63	.370	-.037	0	9
Figuren lernen	4.43	2.05	-.068	-.638	0	8
Fakten lernen	4.24	1.77	-.111	-.426	0	8

Mittels Tabelle 32 werden die Koeffizienten zur Reliabilität der sechs EMS-ähnlichen Subtest dargelegt. „Der Reliabilitätskoeffizient ist ein wichtiger Kennwert zur Beurteilung des Gesamttestes. Er ist das Maß der Genauigkeit, mit der ein Merkmal durch den Test erfasst wird“ (Bühl, 2006, S. 477). Die nachfolgenden Berechnungen der Reliabilitätskoeffizienten werden anhand der internen Konsistenz mittels Cronbach’s Alpha durchgeführt. Angemessene Cronbach α Werte liegen gemäß Kline (1999, zitiert nach Field, 2009, S. 675) bei Intelligenztests bei .8, bei „ability“-Tests bei .7 und bei psychologischen Konstrukten sind auch Werte darunter zu erwarten.

Tabelle 32: Reliabilitätskoeffizienten (Cronbach’s Alpha) der EMS-ähnlichen Untertests

EMS-ähnliche Untertests	Cronbach’s Alpha	Itemanzahl (k)
Quantitative u. formale Probleme	.335	16
Schlauchfiguren	.651	21
Medizinisch-naturw. Grundverständnis	.353	16
Diagramme u. Tabellen	.314	16
Figuren lernen	.637	8
Fakten lernen	.489	8

Die erhobenen Reliabilitätskoeffizienten (Cronbach’s Alpha) weisen bei vier der sechs Skalen, und zwar bei *Quantitative und formale Probleme*, *Medizinisch-naturwissenschaftliches Grundverständnis*, *Diagramme und Tabellen* sowie *Fakten lernen* auf relativ niedrige Konsistenzen hin.

Nachfolgend werden die Koeffizienten zur Reliabilität (Cronbach’s Alpha) der EMS-ähnlichen Untertests in Abhängigkeit der Faktorstufenkombinationen Stereotype Threat und Risikoverhalten dargestellt. Im EMS-ähnlichen Subtest *Quantitative und formale Probleme* kann der höchste Wert der internen Konsistenz nach Cronbach’s Alpha ($\alpha = .432$) in der Faktorstufenkombination Neutral/Sicher beobachtet werden (siehe Tabelle 33).

Tabelle 33: Reliabilitätskoeffizienten (Cronbach’s Alpha) der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnl. Subtest Quantitative und formale Probleme

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.298 (93;16)	.244 (90; 16)	.400 (104; 15)	.098 (86; 15)	.287 (373; 16)
	Ohne Instr.	.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)

Anhand Tabelle 34 erfolgt die Darlegung der Reliabilitätskoeffizienten des EMS-ähnlichen Untertests *Schlauchfiguren* in Abhängigkeit der Faktorstufenkombinationen. Der höchste Wert der internen Konsistenz nach Cronbach's Alpha ($\alpha = .728$) liegt in der Faktorstufenkombination Aktivierung/Raten.

Tabelle 34: Reliabilitätskoeffizienten (Cronbach's Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnl. Subtest Schlauchfiguren

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.728 (93; 21)	.708 (90; 21)	.663 (104; 21)	.596 (86; 21)	.681 (373; 21)
	Ohne Instr.	.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)

Im EMS-ähnlichen Untertest *Medizinisch-naturwissenschaftliches Grundverständnis* ergibt sich der höchste Wert der internen Konsistenz nach Cronbach's Alpha ($\alpha = .476$) in der Faktorstufenkombination Neutral/ohne Instruktion (siehe Tabelle 35).

Tabelle 35: Reliabilitätskoeffizienten (Cronbach's Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnl. Subtest Medizinisch-naturwissenschaftliches Grundverständnis

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.247 (93; 16)	.442 (90; 16)	.398 (104; 16)	.295 (86; 16)	.362 (373; 16)
	Ohne Instr.	.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)

Anhand von Tabelle 36 erfolgt die Darlegung der Koeffizienten zur Reliabilität des EMS-ähnlichen Untertests *Diagramme und Tabellen* in Abhängigkeit der Faktorstufenkombinationen. Der höchste Wert der internen Konsistenz nach Cronbach's Alpha ($\alpha = .434$) liegt in der Faktorstufenkombination Aufklärung/Sicher.

Tabelle 36: Reliabilitätskoeffizienten (Cronbach's Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnl. Subtest Diagramme und Tabellen

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.370 (93, 16)	.341 (90; 15)	.288 (104; 16)	.010 (86; 16)	.289 (373; 16)
	Ohne Instr.	.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)

Im EMS-ähnlichen Untertest *Figuren lernen* ergibt sich der höchste Wert der internen Konsistenz nach Cronbach's Alpha ($\alpha = .752$) in der Faktorstufenkombination Negativ-Aktivierung/Sicher (siehe Tabelle 37).

Tabelle 37: Reliabilitätskoeffizienten (Cronbach's Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnl. Subtest Figuren lernen

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.577 (93; 8)	.682 (90; 8)	.413 (104; 8)	.668 (86; 8)	.595 (373; 8)
	Ohne Instr.	.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)

Tabelle 38 stellt die Koeffizienten zur Reliabilität des EMS-ähnlichen Untertests *Fakten lernen* in Abhängigkeit der Faktorstufenkombinationen dar. Der höchste Wert der internen Konsistenz nach Cronbach's Alpha ($\alpha = .607$) kann in der Faktorstufenkombination Aktivierung/ohne Instruktion beobachtet werden.

Tabelle 38: Reliabilitätskoeffizienten (Cronbach's Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnl. Subtest Fakten lernen

		Stereotype Threat - Bedingungen				Gesamt
		Aktivierung	Neutral	Aufklärung	Negativ-Aktivierung	
Risiko-	Raten	.399 (93;8)	.464 (90;8)	.309 (104;8)	.436 (86;8)	.399 (373;8)
	Ohne Instr.	.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)

9.1.2 Lösungswahrscheinlichkeiten und Itemtrennschärfen in den EMS-ähnlichen Subtests

In diesem Abschnitt werden die relativen Lösungshäufigkeiten als Schätzung für die Lösungswahrscheinlichkeiten der vorgegebenen Aufgaben in den sechs EMS-ähnlichen Untertests und die korrigierten Itemtrennschärfen dargelegt.

Die Lösungswahrscheinlichkeit eines Items ist als Itemschwierigkeit quantifizierbar. Der Schwierigkeitsindex eines Items (Wertebereich 0-1) charakterisiert den Anteil der Probanden, die das Item richtig lösen und je niedriger dieser Wert ist, desto schwerer fällt es den Probanden, das Item zu lösen. Der Schwierigkeitsindex eines Items sollte im mittleren Bereich liegen (0.2 – 0.8) um Personenunterschiede aufzeigen zu können, dies gelingt bei sehr leichten respektive sehr schwierigen Items nicht (Bortz & Döring, 2005).

Anhand der Itemtrennschärfe kann entnommen werden, „wie gut ein einzelnes Item das Gesamtergebnis eines Tests repräsentiert“ (Bortz & Döring, 2005, S. 218).

Die Itemtrennschärfe entspricht dem korrelativen Zusammenhang der Beantwortung des Items mit dem gesamten Testwert und dementsprechend sind hinreichend hohe Trennschärfen (≥ 0.30) wünschenswert. Itemtrennschärfen hängen auch mit der Itemschwierigkeit zusammen, wobei die Trennschärfe eines Items umso niedriger ist, je schwieriger das Item ist (Bortz & Döring, 2005).

Die Tabelle 39 zeigt die Lösungswahrscheinlichkeiten und die korrigierten Itemtrennschärfen des EMS-ähnlichen Untertests *Quantitative und formale Probleme*.

Tabelle 39: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Quantitative und formale Probleme

Quantitative und formale Probleme				
Items		N = 1093		
		M	SD	korrigierte Itemtrennschärfe
1	Quant 1	0.54	0.50	.052
2	Quant 2	0.47	0.50	.038
3	Quant 3	0.26	0.44	.085
4	Quant 4	0.27	0.45	.114
5	Quant 73	0.51	0.50	.142
6	Quant 74	0.27	0.44	.216

	Items	M	SD	korrigierte Itemtrennschärfe
7	Quant 75	0.21	0.41	.218
8	Quant 76	0.11	0.31	.172
9	Quant 5	0.16	0.37	.144
10	Quant 6	0.04	0.20	.058
11	Quant 7	0.02	0.16	.084
12	Quant 8	0.02	0.15	.039
13	Quant 93	0.03	0.18	.019
14	Quant 94	0.03	0.16	.079
15	Quant 95	0.02	0.15	.068
16	Quant 96	0.02	0.14	.068

Die relativen Lösungswahrscheinlichkeiten in diesem Untertest deuten auf Items im mittelschwierigen bis schwierigen Bereich hin, wobei die Zunahme der Itemschwierigkeiten ab etwa der zweiten Hälfte der Aufgaben vermutlich auf die kurze Testbearbeitungszeit (10 Minuten) zurückgeht. Die zum Teil recht niedrigen Trennschärfekoeffizienten weisen zudem darauf hin, dass einzelne Items nicht dasselbe wie der Gesamttest erfassen (vgl. Bortz & Döring, 2005).

Anhand der Abbildung 18 wird ersichtlich, dass ab der 9. Aufgabe eine deutliche Verringerung der Lösungshäufigkeiten angenommen werden kann.

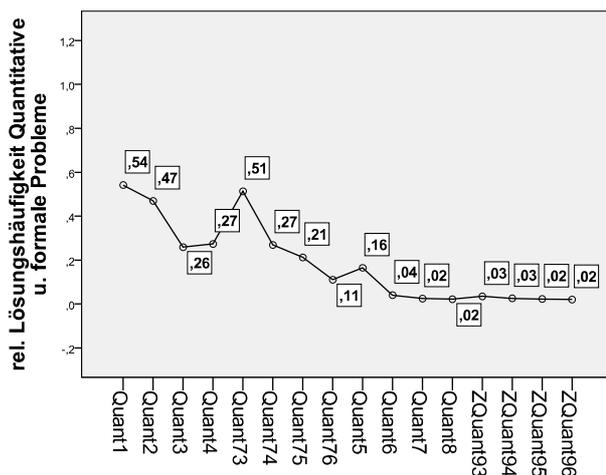


Abbildung 18: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Quantitative und formale Probleme

Mittels Tabelle 40 werden die Lösungswahrscheinlichkeiten und die korrigierten Itemtrennschärfen im EMS-ähnlichen Subtest *Schlauchfiguren* ersichtlich.

Tabelle 40: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Schlauchfiguren

Schlauchfiguren				
Items		N = 1093		
		M	SD	korrigierte Itemtrennschärfe
1	Schlauch 1	0.74	0.44	.290
2	Schlauch 2	0.67	0.47	.251
3	Schlauch 3	0.56	0.50	.230
4	Schlauch 4	0.72	0.45	.371
5	Schlauch 5	0.32	0.47	.124
6	Schlauch 6	0.65	0.48	.305
7	Schlauch 7	0.50	0.50	.335
8	Schlauch 8	0.53	0.50	.283
9	Schlauch 9	0.34	0.47	.207
10	Schlauch 10	0.63	0.48	.263
11	Schlauch 11	0.59	0.49	.180
12	Schlauch 12	0.71	0.45	.333
13	Schlauch 13	0.68	0.47	.400
14	Schlauch 14	0.31	0.46	.239
15	Schlauch 15	0.25	0.43	.188
16	Schlauch 16	0.37	0.48	.177
17	Schlauch 17	0.37	0.48	.155
18	Schlauch 18	0.27	0.45	.109
19	Schlauch 19	0.61	0.49	.155
20	Schlauch 20	0.09	0.29	.085
21	Schlauch 21	0.28	0.45	.102

Die Lösungswahrscheinlichkeiten liegen mit einer Ausnahme im mittleren Bereich, was auf mittelschwierige Items in diesem EMS-ähnlichen Untertest hindeutet. Fünf Items können als trennscharf (≥ 0.30) bezeichnet werden.

Anhand von Abbildung 19 wird die relativ niedrige Lösungswahrscheinlichkeit beim vorletzten (20.) Item ersichtlich.

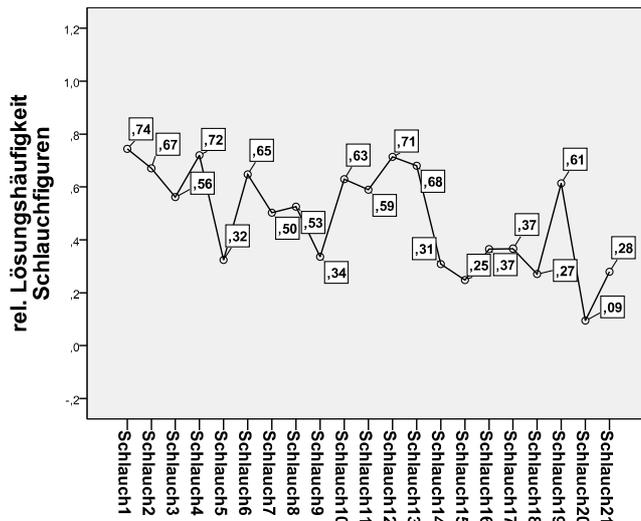


Abbildung 19: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Schlauchfiguren

Die Lösungshäufigkeiten und die korrigierten Itemtrennschärfen des EMS-ähnlichen Untertests *Medizinisch-naturwissenschaftliches Grundverständnis* werden in Tabelle 41 dargelegt.

Tabelle 41: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest *Medizinisch-naturwissenschaftliches Grundverständnis*

Medizinisch-naturwissenschaftl. Grundverständnis				
Items		N = 1093		
		M	SD	korrigierte Itemtrennschärfe
1	Natur 32	0.56	0.50	.120
2	Natur 33	0.39	0.49	.153
3	Natur 34	0.38	0.49	.106
4	Natur 35	0.19	0.40	.134
5	Natur 26	0.29	0.45	.195
6	Natur 27	0.30	0.46	.113
7	Natur 28	0.21	0.41	.146
8	Natur 29	0.16	0.37	.148
9	Natur 36	0.07	0.25	.100
10	Natur 37	0.05	0.22	.029
11	Natur 38	0.03	0.17	.013
12	Natur 39	0.04	0.19	.079
13	Natur 44	0.04	0.19	.102
14	Natur 45	0.03	0.18	.036
15	Natur 46	0.04	0.20	.090
16	Natur 47	0.03	0.16	.066

In diesem Untertest werden überwiegend recht niedrige Itemtrennschärfen beobachtet. Die relativen Lösungswahrscheinlichkeiten weisen auf Items im mittelschwierigen bis schwierigen Bereich hin. Anhand des Verlaufsdiagramms zur Lösungshäufigkeit (siehe Abbildung 20) kann angenommen werden, dass die Zunahme der Itemschwierigkeiten auf die geringe Bearbeitungszeit (10 Minuten) rückführbar ist.

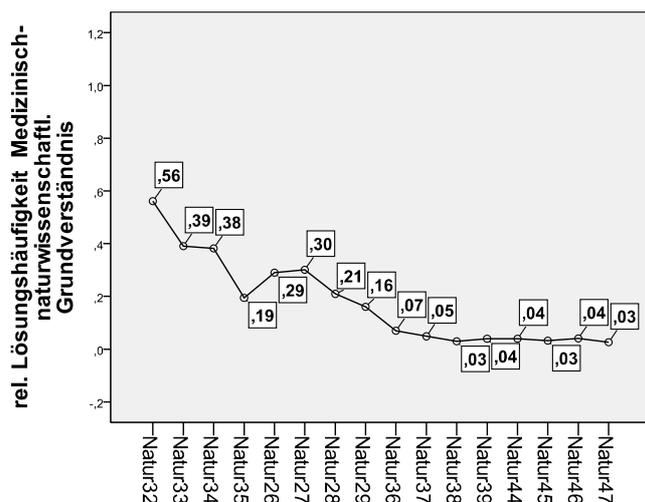


Abbildung 20: Relative Lösungshäufigkeit im EMS-ähnlichen Untertest Medizinisch-naturwissenschaftliches Grundverständnis

Anhand von Tabelle 42 werden die Lösungswahrscheinlichkeiten und die korrigierten Itemtrennschärfen des EMS-ähnlichen Subtests *Diagramme und Tabellen* dargelegt.

Tabelle 42: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest *Diagramme und Tabellen*

Diagramme und Tabellen				
Items		N = 1093		
		M	SD	korrigierte Itemtrennschärfe
1	Diagr 71	0.58	0.49	.104
2	Diagr 72	0.43	0.50	.095
3	Diagr 73	0.59	0.49	.050
4	Diagr 74	0.25	0.43	.109
5	Diagr 161	0.27	0.44	.098
6	Diagr 162	0.27	0.45	.172
7	Diagr 163	0.13	0.33	.154

	Items	<i>M</i>	<i>SD</i>	korrigierte Itemtrennschärfe
8	Diagr 164	0.16	0.37	.103
9	Diagr 75	0.09	0.28	.091
10	Diagr 76	0.05	0.23	.091
11	Diagr 77	0.03	0.17	.071
12	Diagr 78	0.03	0.18	.032
13	Diagr 181	0.02	0.15	.071
14	Diagr 182	0.05	0.22	.083
15	Diagr 183	0.04	0.19	.124
16	Diagr 184	0.03	0.17	.090

In diesem Untertest werden niedrige Itemtrennschärfen beobachtet und die relativen Lösungswahrscheinlichkeiten deuten auf Items im mittelschwierigen bis schwierigen Bereich hin. Für die Verringerung der Lösungshäufigkeit ab der zweiten Hälfte der Aufgaben spielt vermutlich die Bearbeitungszeit (10 Minuten) eine bedeutende Rolle (siehe Abbildung 21).

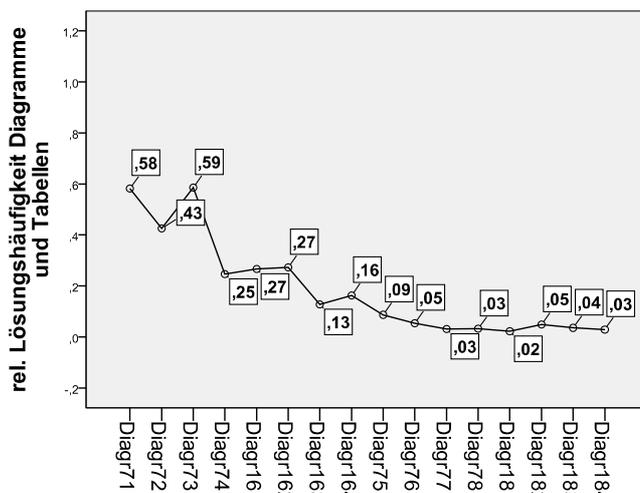


Abbildung 21: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Diagramme und Tabellen

Die Lösungswahrscheinlichkeiten und die korrigierten Itemtrennschärfen des EMS-ähnlichen Subtests *Figuren lernen* sind Tabelle 43 zu entnehmen.

Tabelle 43: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Figuren lernen

Figuren lernen				
Items		N = 1093		
		M	SD	korrigierte Itemtrennschärfe
1	Reprod. Figuren 47	0.59	0.49	.283
2	Reprod. Figuren 48	0.57	0.50	.316
3	Reprod. Figuren 49	0.70	0.46	.360
4	Reprod. Figuren 50	0.43	0.50	.339
5	Reprod. Figuren 51	0.38	0.49	.330
6	Reprod. Figuren 52	0.70	0.46	.380
7	Reprod. Figuren 53	0.42	0.49	.318
8	Reprod. Figuren 54	0.63	0.48	.322

Die Lösungswahrscheinlichkeiten weisen auf Items im mittelschwierigen Bereich hin und die hinreichend hohen Trennschärfekoeffizienten (außer bei Item Nr. 47) deuten darauf hin, dass die einzelnen Items dasselbe wie der Gesamttest erfassen (vgl. Bortz & Döring, 2005). Anhand von Abbildung 22 ist der Verlauf der relativen Lösungshäufigkeit ersichtlich.

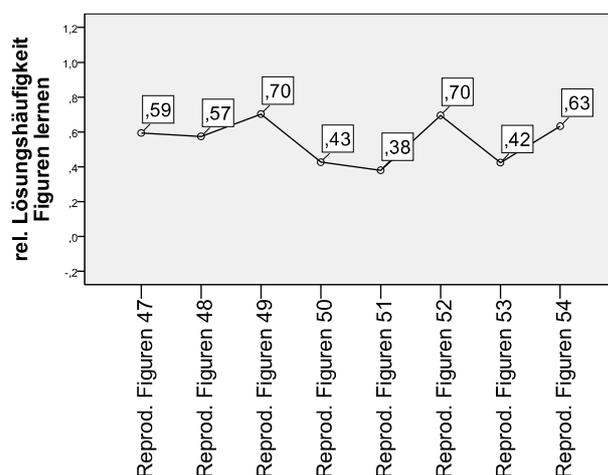


Abbildung 22: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Figuren lernen

Die Tabelle 44 zeigt die Lösungswahrscheinlichkeiten und die korrigierten Itemtrennschärfen des EMS-ähnlichen Untertests *Fakten lernen*.

Tabelle 44: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Fakten lernen

Fakten lernen				
Items		N = 1093		
		<i>M</i>	<i>SD</i>	korrigierte Itemtrennschärfe
1	Reprod. Fakten 55	0.65	0.48	.245
2	Reprod. Fakten 56	0.43	0.50	.218
3	Reprod. Fakten 57	0.35	0.48	.120
4	Reprod. Fakten 58	0.52	0.50	.300
5	Reprod. Fakten 59	0.85	0.36	.228
6	Reprod. Fakten 60	0.39	0.49	.265
7	Reprod. Fakten 61	0.65	0.48	.241
8	Reprod. Fakten 62	0.40	0.49	.166

In diesem Untertest werden überwiegend niedrige Itemtrennschärfen beobachtet und die Lösungswahrscheinlichkeiten liegen mehrheitlich im mittleren Bereich. Die Lösungswahrscheinlichkeit der fünften Aufgabe weist auf ein eher leicht lösbares Item hin (siehe Abbildung 23).

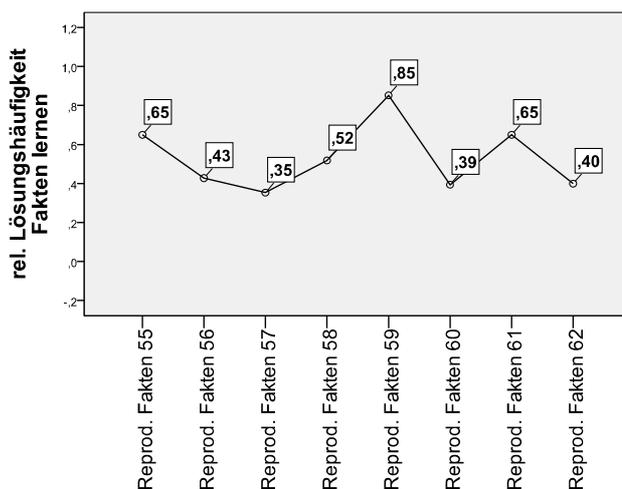


Abbildung 23: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Fakten lernen

9.1.3 Selbstwirksamkeitserwartung

Im Folgenden werden die Itemstatistiken (siehe Tabelle 45) sowie die Kennwerte und Koeffizienten zur Selbstwirksamkeitserwartung (SWE) dargelegt.

Tabelle 45: Itemstatistik Selbstwirksamkeitserwartung (N = 1034 gültige Fälle)

	<i>M</i>	<i>SD</i>	korrigierte Itemtrennschärfe
SWE1	3.13	0.61	.385
SWE2	3.16	0.65	.429
SWE3	2.90	0.71	.398
SWE4	2.57	0.75	.447
SWE5	2.91	0.66	.453
SWE6	2.71	0.79	.493
SWE7	3.16	0.76	.417
SWE8	2.98	0.78	.435
SWE9	2.76	0.69	.499
SWE10	2.97	0.66	.458

Die einzelnen Trennschärfekoeffizienten deuten auf eine zufrieden stellende Item-Skalenkorrelation hin und weisen zusammen mit dem Cronbach Alpha Wert mit $\alpha = .775$ auf eine konsistente Skala hin.

Auf Grund der Konsistenz der Items kann unter Berücksichtigung der Verrechnungsvorschriften nach Schwarzer (2009) ein ungewichteter additiver Score errechnet werden, wobei die Meanscore-Bildung nur bei jenen 1071 UntersuchungsteilnehmerInnen durchzuführen war, die zumindest sieben oder mehr Items der *Skala zur Allgemeinen Selbstwirksamkeitserwartung* (Jerusalem & Schwarzer, 1999) beantwortet haben.

Anhand von Tabelle 46 sind die Kennwerte zur Selbstwirksamkeitserwartung ersichtlich.

Tabelle 46: Deskriptivstatistik Selbstwirksamkeitserwartung

SWE – Meanscore		
N	Gültig	1071
	Fehlend	22
M		2.93
SD		0.41
Schiefe		-.204
Kurtosis		.335
Min.		1.40
Max.		4.00

Die Normalverteilung des Scores kann angenommen werden, die Werte der Schiefe und Kurtosis weichen nicht wesentlich von Null ab. Abbildung 24 zeigt die Häufigkeitsverteilung in der Selbstwirksamkeitserwartung.

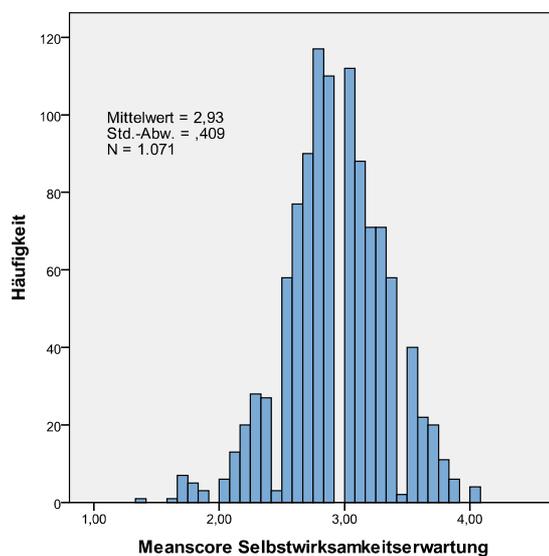


Abbildung 24: Histogramm zum Meanscore SWE

9.1.4 Game of Dice Task

Anhand von Tabelle 47 werden die Testkennwerte zum Game of Dice Task (GDT) (Brand et al., 2005) dargelegt. Der Reliabilitätskoeffizient (Cronbach's Alpha) weist mit $\alpha = .891$ auf eine hinreichende Konsistenz hin.

Tabelle 47: Deskriptivstatistik Game of Dice Task

Nettoscore GDT		
N	Gültig	471
	Fehlend	622
M		4.31
SD		9.71
Schiefe		-.516
Kurtosis		-.645
Min.		-18
Max.		18

Aufgrund der Werte der Schiefe und Kurtosis kann eine annähernde Normalverteilung der Daten angenommen werden. Anhand von Abbildung 25 wird die Häufigkeitsverteilung im Game of Dice Task dargestellt.

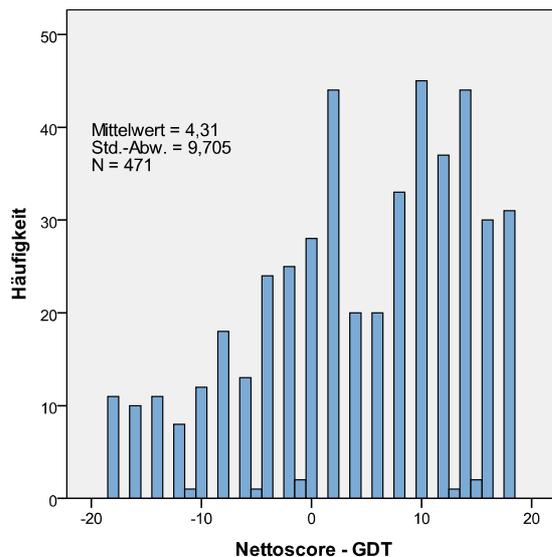


Abbildung 25: Histogramm zum Nettoscore-GDT

9.2 Überprüfung der Wirksamkeit der Stereotype Threat-Untersuchungsbedingungen

Die Überprüfung der Wirksamkeit der vier Stereotype Threat-Untersuchungsbedingungen:

- Aktivierung: Frauen erzielen bessere Leistungen im Gedächtnis, während Männer bessere Leistungen in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen und im naturwissenschaftlichen Wissen aufweisen (vgl. Smith & White, 2002),
- Neutral: Es gibt keine geschlechtsspezifischen Unterschiede in den Leistungen zwischen Männern und Frauen (vgl. Spencer et al., 1999),
- Aufklärung: über Stereotype Threat wird aufgeklärt (vgl. Johns et al., 2005),
- Negativ-Aktivierung: Männern weisen bessere Leistungen im Gedächtnis auf, während Frauen bessere Leistungen in der räumlichen Orientierung, im Arbeiten mit Zahlen, in der Interpretation von Diagrammen und Tabellen und im naturwissenschaftlichen Wissen erzielen,

erfolgte anhand des Wissenstests (siehe Unterpunkt „Erhebungsinstrumente“), der auf Basis des „Beliefs About the Cognitive Abilities of Females and Males Questionnaire“ von Halpern und Tan (2001) konstruiert wurde. Bei Halpern und Tan (2001) wurde festgelegt, dass die Wahrscheinlichkeitseinschätzungen bei dem jeweiligen Item, dass es sich um eine Frau bzw. Mann handelt, innerhalb des Mittelwertbereiches von 45,0 bis 55,0 Prozent unwesentlich von 50 Prozent abweichen und demgemäß gleich wahrscheinlich mit Frauen und Männern assoziiert werden.

In der vorliegenden Untersuchung entsprechen in den Untersuchungsbedingungen *Neutral* und *Aufklärung* Prozentangaben innerhalb des Bereiches von 45,0 bis 55,0 Prozent einer wirksamen Manipulation, während in den Untersuchungsbedingungen *Aktivierung* und *Negativ-Aktivierung* Prozentangaben außerhalb dieses Bereiches effektive Manipulationen darstellen.

Anhand von Tabelle 48 wird die Häufigkeitsverteilung der Items zur geschlechtsspezifischen Fähigkeitszuschreibung⁴ in Abhängigkeit von der Wirksamkeit der vorgegebenen Instruktionen dargelegt.

⁴ Die geschlechtsspezifische und persönliche Fähigkeitszuschreibung zur Theory of Mind wird von Katrin Anzirk hinsichtlich ihres Schwerpunkts untersucht.

Tabelle 48: Häufigkeitsverteilung der Items zur geschlechtsspezifischen Fähigkeitszuschreibung in Abhängigkeit von der Wirksamkeit der Instruktionen

Geschlechtsspezifische Fähigkeitszuschreibung							
Wirksamkeit der Instruktionen		Quant. u. formale Probleme	Schlauchfiguren	Med.-naturw. Grundv.	Diagr. u. Tabellen	Figuren lernen	Fakten lernen
		Häufigkeit	Häufigkeit	Häufigkeit	Häufigkeit	Häufigkeit	Häufigkeit
Gültig	nein	339	314	334	320	288	307
	ja	590	651	571	637	646	580
	entgegen-gesetzt	151	115	175	123	150	195
	Gesamt	1080	1080	1080	1080	1084	1082
Fehlend	System	13	13	13	13	9	11
Gesamt		1093	1093	1093	1093	1093	1093

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt bei allen sechs Items ein signifikantes Ergebnis (siehe Tabelle 49). Es kann ein Verteilungsunterschied in den sechs Items zur geschlechtsspezifischen Fähigkeitszuschreibung angenommen werden

Tabelle 49: Chi-Quadrat-Tests Items zur geschlechtsspezifischen Fähigkeitszuschreibung

Geschlechtsspezifische Fähigkeitszuschreibung						
	Quantitative u. formale Probleme	Schlauchfiguren	Med.-naturw. Grundv.	Diagramme u. Tabellen	Figuren lernen	Fakten lernen
Chi-Quadrat	269.506 ^c	407.839 ^c	220.617 ^c	373.606 ^c	362.753 ^a	217.466 ^b
df	2	2	2	2	2	2
Asympt. Sign.	≤ .0001	≤ .0001	≤ .0001	≤ .0001	≤ .0001	≤ .0001

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 361,3.

b. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 360,7.

c. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 360,0.

Nachfolgend wird anhand der sechs Items zur geschlechtsspezifischen Fähigkeitszuschreibung die Wirksamkeit der vier Stereotype Threat-Untersuchungsbedingungen überprüft.

Auf Grundlage der Kontingenztafel (siehe Tabelle 50) ist die Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Quantitative und formale Probleme* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen ersichtlich.

Tabelle 50: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Quantitative und formale Probleme

Stereotype Threat * Quantitative und formale Probleme Kreuztabelle						
			Wirksamkeit der Instruktion bei Quant. u. formale Probleme			Gesamt
			nein	ja	entgegengesetzt	
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	60	207	22	289
		Erwartete Anz.	90.7	157,9	40.4	289.0
		% innerhalb	20,8%	71,6%	7,6%	100,0%
		Stand. Residuen	-3.2	3.9	-2.9	
	Neutral	Anzahl	86	142	26	254
		Erwartete Anz.	79.7	138.8	35.5	254.0
		% innerhalb	33,9%	55,9%	10,2%	100,0%
		Stand. Residuen	.7	.3	-1.6	
	Aufklärung	Anzahl	141	114	22	277
		Erwartete Anz.	86.9	151.3	38.7	277.0
		% innerhalb	50,9%	41,2%	7,9%	100,0%
		Stand. Residuen	5.8	-3.0	-2.7	
	Negativ-Aktivierung	Anzahl	52	127	81	260
		Erwartete Anz.	81.6	142.0	36.4	260.0
		% innerhalb	20,0%	48,8%	31,2%	100,0%
		Stand. Residuen	-3.3	-1.3	7.4	
Gesamt	Anzahl	339	590	151	1080	
	Erwartete Anz.	339.0	590.0	151.0	1080.0	
	% innerhalb	31,4%	54,6%	14,0%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 154.393$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur

geschlechtsspezifischen Fähigkeitszuschreibung *Quantitative und formale Probleme* in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen beobachtet werden.

Die Darstellung der Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Schlauchfiguren* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen, erfolgt in der Tabelle 51.

Tabelle 51: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Schlauchfiguren

Stereotype Threat * Schlauchfiguren Kreuztabelle						
		Wirksamkeit der Instruktion bei Schlauchfiguren			Gesamt	
		nein	ja	entgegengesetzt		
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	70	210	9	289
		Erwartete Anz.	84.0	174.2	30.8	289.0
		% innerhalb	24,2%	72,7%	3,1%	100,0%
		Stand. Residuen	-1.5	2.7	-3.9	
	Neutral	Anzahl	77	149	28	254
		Erwartete Anz.	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 Anz.	80.5	167.0	29.5	277.0
		% innerhalb	36,8%	53,1%	10,1%	100,0%
		Stand. Residuen	2.4	-1.5	-.3	
	Negativ-Aktivierung	Anzahl	65	145	50	260
		Erwartete Anz.	75.6	156.7	27.7	260.0
		% innerhalb	25,0%	55,8%	19,2%	100,0%
		Stand. Residuen	-1.2	-.9	4.2	
Gesamt	Anzahl	314	651	115	1080	
	Erwartete Anz.	314.0	651.0	115.0	1080.0	
	% innerhalb	29,1%	60,3%	10,6%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 53.913$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur

geschlechtsspezifischen Fähigkeitszuschreibung *Schlauchfiguren* in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen aufgezeigt werden.

Die Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Medizinisch-naturwissenschaftliches Grundverständnis* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen, wird anhand von Tabelle 52 ersichtlich.

Tabelle 52: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Medizinisch-naturwissenschaftliches Grundverständnis

Stereotype Threat* Med.-naturw. Grundverständnis Kreuztabelle						
			Wirksamkeit der Instruktion bei Medizinisch-naturw. Grundverständnis			Gesamt
			nein	ja	entgegengesetzt	
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	100	149	41	290
		Erwartete Anz.	89.7	153.3	47.0	290.0
		% innerhalb	34,5%	51,4%	14,1%	100,0%
		Stand. Residuen	1.1	-.3	-.9	
	Neutral	Anzahl	76	155	23	254
		Erwartete Anz.	78.6	134.3	41.2	254.0
		% innerhalb	29,9%	61,0%	9,1%	100,0%
		Stand. Residuen	-.3	1.8	-2.8	
	Aufklärung	Anzahl	87	150	40	277
		Erwartete Anz.	85.7	146.5	44.9	277.0
		% innerhalb	31,4%	54,2%	14,4%	100,0%
		Stand. Residuen	.1	.3	-.7	
	Negativ-Aktivierung	Anzahl	71	117	71	259
		Erwartete Anz.	80.1	136.9	42.0	259.0
		% innerhalb	27,4%	45,2%	27,4%	100,0%
		Stand. Residuen	-1.0	-1.7	4.5	
Gesamt	Anzahl	334	571	175	1080	
	Erwartete Anz.	334.0	571.0	175.0	1080.0	
	% innerhalb	30,9%	52,9%	16,2%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 38.017$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Medizinisch-naturwissenschaftliches*

Grundverständnis in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen beobachtet werden.

Anhand der Kontingenztafel (Tabelle 53) wird die Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Diagramme und Tabellen* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen dargestellt.

Tabelle 53: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Diagramme und Tabellen

Stereotype Threat * Diagramme und Tabellen Kreuztabelle						
		Wirksamkeit der Instruktion bei Diagramme und Tabellen			Gesamt	
		nein	ja	entgegengesetzt		
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	52	223	14	289
		Erwartete Anz.	85.6	170.5	32.9	289.0
		% innerhalb	18,0%	77,2%	4,8%	100,0%
		Stand. Residuen	-3.6	4.0	-3.3	
	Neutral	Anzahl	95	139	20	254
		Erwartete Anz.	75.3	149.8	28.9	254.0
		% innerhalb	37,4%	54,7%	7,9%	100,0%
		Stand. Residuen	2.3	-.9	-1.7	
	Aufklärung	Anzahl	125	123	29	277
		Erwartete Anz.	82.1	163.4	31.5	277.0
		% innerhalb	45,1%	44,4%	10,5%	100,0%
		Stand. Residuen	4.7	-3.2	-.5	
	Negativ-Aktivierung	Anzahl	48	152	60	260
		Erwartete Anz.	77.0	153.4	29.6	260.0
		% innerhalb	18,5%	58,5%	23,1%	100,0%
		Stand. Residuen	-3.3	-.1	5.6	
Gesamt	Anzahl	320	637	123	1080	
	Erwartete Anz.	320.0	637.0	123.0	1080.0	
	% innerhalb	29,6%	59,0%	11,4%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 123.767$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur

geschlechtsspezifischen *Diagramme und Tabellen* in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen festgestellt werden.

Die Kontingenztafel (Tabelle 54) lässt die Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Figuren lernen* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen erkennen.

Tabelle 54: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Figuren lernen

Stereotype Threat * Figuren lernen Kreuztabelle						
			Wirksamkeit der Instruktion bei Figuren lernen			Gesamt
			nein	ja	entgegengesetzt	
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	115	124	50	289
		Erwartete Anz.	76.8	172.2	40.0	289.0
		% innerhalb	39,8%	42,9%	17,3%	100,0%
		Stand. Residuen	4.4	-3.7	1.6	
	Neutral	Anzahl	37	205	13	255
		Erwartete Anz.	67.7	152.0	35.3	255.0
		% innerhalb	14,5%	80,4%	5,1%	100,0%
		Stand. Residuen	-3.7	4.3	-3.8	
	Aufklärung	Anzahl	33	216	30	279
		Erwartete Anz.	74.1	166.3	38.6	279.0
		% innerhalb	11,8%	77,4%	10,8%	100,0%
		Stand. Residuen	-4.8	3.9	-1.4	
	Negativ-Aktivierung	Anzahl	103	101	57	261
		Erwartete Anz.	69.3	155.5	36.1	261.0
		% innerhalb	39,5%	38,7%	21,8%	100,0%
		Stand. Residuen	4.0	-4.4	3.5	
Gesamt	Anzahl	288	646	150	1084	
	Erwartete Anz.	288.0	646.0	150.0	1084.0	
	% innerhalb	26,6%	59,6%	13,8%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 168.720$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Figuren lernen* in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen beobachtet werden.

Auf Grundlage der Kontingenztafel (Tabelle 55) wird die Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Fakten lernen* in Abhängigkeit von den vier Stereotype Threat-Untersuchungsbedingungen dargestellt.

Tabelle 55: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Fakten lernen

Stereotype Threat * Fakten lernen Kreuztabelle						
			Wirksamkeit der Instruktion bei Fakten lernen			Gesamt
			nein	ja	entgegengesetzt	
Stereotype Threat - Bedingungen	Aktivierung	Anzahl	48	205	37	290
		Erwartete Anz.	82.3	155.5	52.3	290.0
		% innerhalb	16,6%	70,7%	12,8%	100,0%
		Stand. Residuen	-3.8	4.0	-2.1	
	Neutral	Anzahl	84	132	38	254
		Erwartete Anz.	72.1	136.2	45.8	254.0
		% innerhalb	33,1%	52,0%	15,0%	100,0%
		Stand. Residuen	1.4	-.4	-1.1	
	Aufklärung	Anzahl	127	104	47	278
		Erwartete Anz.	78.9	149.0	50.1	278.0
		% innerhalb	45,7%	37,4%	16,9%	100,0%
		Stand. Residuen	5.4	-3.7	-.4	
	Negativ-Aktivierung	Anzahl	48	139	73	260
		Erwartete Anz.	73.8	139.4	46.9	260.0
		% innerhalb	18,5%	53,5%	28,1%	100,0%
		Stand. Residuen	-3.0	.0	3.8	
Gesamt	Anzahl	307	580	195	1082	
	Erwartete Anz.	307.0	580.0	195.0	1082.0	
	% innerhalb	28,4%	53,6%	18,0%	100,0%	

Die Berechnung des Vergleichs der empirischen vs. der theoretisch zu erwartenden Verteilung ergibt mit $\chi^2(6) = 104.697$, $p \leq .0001$ ein signifikantes Ergebnis. Es kann ein Verteilungsunterschied in der Wirksamkeit der Instruktionen bei dem Item zur geschlechtsspezifischen Fähigkeitszuschreibung *Fakten lernen* in Abhängigkeit von den Stereotype Threat-Untersuchungsbedingungen aufgezeigt werden.

Die Ergebnisse der durchgeführten Chi-Quadrat-Anpassungstests und die Mittelwerte der Prozentangaben zu den sechs Items der geschlechtsspezifischen

Fähigkeitszuschreibung (siehe Anhang) untermauern die Wirksamkeit der vier Stereotype Threat-Untersuchungsbedingungen.

Nachfolgend wird der Einfluss von Geschlecht und Stereotype Threat auf die persönlichen Fähigkeitszuschreibungen analysiert. Es wird eine zweifaktorielle multivariate Varianzanalyse (MANOVA) mit den sechs Items zur persönlichen Fähigkeitszuschreibung der EMS Untertests als abhängige Variablen und den Zwischensubjektfaktoren Geschlecht und Stereotype Threat durchgeführt.

Die Levene-Tests (siehe Tabelle 56) ergeben, dass bei den persönlichen Fähigkeitszuschreibungen zu *Figuren lernen* ($p = .002$) und *Schlauchfiguren* ($p \leq .0001$) die Homogenität der Varianzen nicht angenommen werden kann und der Box's-M Test ergibt mit $p \leq .0001$. ein signifikantes Ergebnis, womit die Homogenität der Varianz-Kovarianz Matrizen nicht gegeben ist. Aufgrund der Robustheit der Varianzanalyse kann jedoch die Hypothesenprüfung trotzdem durchgeführt werden.

Tabelle 56: Levene's Test, persönliche Fähigkeitszuschreibungen

Levene-Test auf Gleichheit der Fehlervarianzen		
Persönliche Fähigkeitszuschreibung	<i>F</i> (7,1057)	Sig.
Figuren lernen	3.22	.002
Fakten lernen	.99	.440
Med.-naturw. Grundverständnis	1.19	.304
Diagramme und Tabellen	2.01	.051
Quantitative und formale Probleme	1.14	.333
Schlauchfiguren	8.62	$\leq .0001$

Die Tabelle 57 stellt die Ergebnisse der multivariaten Varianzanalyse in Abhängigkeit vom Geschlecht dar. Es ist ersichtlich, dass bei den persönlichen Fähigkeitszuschreibungen zu *Figuren lernen* ($p = .031$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .022$), *Diagramme und Tabellen* ($p \leq .0001$), *Quantitative und formale Probleme* ($p \leq .0001$) und *Schlauchfiguren* ($p \leq .0001$) signifikante Unterschiede zwischen den männlichen und weiblichen Testpersonen angenommen werden können. Nur bei der persönlichen Fähigkeitszuschreibung *Fakten lernen* kann kein signifikanter Unterschied in Abhängigkeit vom Geschlecht beobachtet werden.

Tabelle 57: Tests der Zwischensubjekteffekte, Quelle: Geschlecht, AV: persönl. Fähigkeitszuschreibung

Persönliche Fähigkeitszuschreibung	Quadratsumme vom Typ III	Mittel der Quadrate	F(1,1057)	Sig.
Figuren lernen	4.50	4.50	4.67	.031
Fakten lernen	2.74	2.74	1.75	.186
Med.-naturw. Grundverständnis	10.38	10.38	5.28	.022
Diagramme und Tabellen	182.05	182.05	73.55	≤ .0001
Quantitative u. formale Probleme	238.59	238.59	78.30	≤ .0001
Schlauchfiguren	127.05	127.05	48.40	≤ .0001

Anhand von Tabelle 58 erfolgt die Darstellung der Deskriptivstatistik (die gesamte Tabelle ist im Anhang zu finden) zu den persönlichen Fähigkeitszuschreibungen in Abhängigkeit vom Geschlecht. Die Mittelwerte verdeutlichen, dass sich männliche Testpersonen bei den Items zu *Figuren lernen*, *Medizinisch-naturwissenschaftliches Grundverständnis*, *Diagramme und Tabellen*, *Quantitative und formale Probleme* sowie *Schlauchfiguren* höhere Fähigkeiten zuschreiben als die weiblichen Untersuchungsteilnehmer.

Tabelle 58: Deskriptivstatistik persönliche Fähigkeitszuschreibungen in Abhängigkeit vom Geschlecht (Frauen: N = 670, Männer: N = 395)

Persönliche Fähigkeitszuschreibung		M	SD	
Geschlecht	Frauen	Figuren lernen	6.50	1.04
		Fakten lernen	5.05	1.23
		Med.-naturw. Grundverständnis	4.86	1.39
		Diagramme und Tabellen	4.22	1.60
		Quantitative u. formale	4.04	1.79
		Schlauchfiguren	5.16	1.73
	Männer	Figuren lernen	6.63	0.87
		Fakten lernen	4.93	1.30
		Med.-naturw. Grundverständnis	5.09	1.43
		Diagramme und Tabellen	5.09	1.57
		Quantitative u. formale	5.05	1.68
		Schlauchfiguren	5.89	1.40

Wie Tabelle 59 zeigt, können keine signifikanten Unterschiede ($p \geq .155$) in den persönlichen Fähigkeitszuschreibungen zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) angenommen werden.

Tabelle 59: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat, AV: persönl. Fähigkeitszuschreibung

Persönliche Fähigkeitszuschreibung	Quadratsumme vom Typ III	Mittel der Quadrate	F(3,1057)	Sig.
Figuren lernen	.98	.33	.34	.797
Fakten lernen	8.21	2.74	1.75	.155
Med.-naturw. Grundverständnis	5.73	1.91	.97	.406
Diagramme und Tabellen	11.83	3.94	1.59	.189
Quantitative u. formale Probleme	12.68	4.23	1.39	.245
Schlauchfiguren	1.73	.58	.22	.882

Eine signifikante Wechselwirkung zwischen Geschlecht * Stereotype Threat-Bedingungen kann bei den persönlichen Fähigkeitszuschreibungen *Diagramme und Tabellen* ($p \leq .0001$) beobachtet werden (siehe Tabelle 60).

Tabelle 60: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat, AV: persönl. Fähigkeitszuschreibung

Persönliche Fähigkeitszuschreibung	Quadratsumme vom Typ III	Mittel der Quadrate	F(3,1057)	Sig.
Figuren lernen	2.95	.98	1.02	.383
Fakten lernen	5.63	1.88	1.20	.309
Med.-naturw. Grundverständnis	9.20	3.07	1.56	.198
Diagramme und Tabellen	50.64	16.88	6.82	$\leq .0001$
Quantitative u. formale Probleme	21.89	7.30	2.40	.067
Schlauchfiguren	11.68	3.89	1.48	.218

Im Folgenden wird diese signifikante Wechselwirkung dargestellt (siehe Abbildung 26).

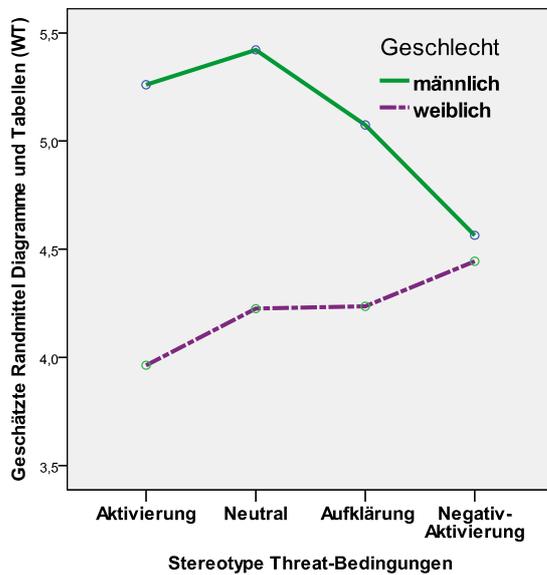


Abbildung 26: Wechselwirkung Geschlecht * Stereotype Threat bei persönlicher Fähigkeitszuschreibung Diagramme und Tabellen

Die Abbildung 26 verdeutlicht, dass sich die weiblichen Testpersonen bei dem Item *Diagramme und Tabellen* in der Bedingung Negativ-Aktivierung die höchsten Fähigkeiten zuschreiben, während sie sich in der Aktivierungs-Bedingung die geringsten Fähigkeiten zuweisen. Bei den männlichen Untersuchungsteilnehmern zeigt sich, dass sie sich bei dem Item *Diagramme und Tabellen* in der Bedingung Neutral die höchsten Fähigkeiten zuschreiben, wohingegen die Negativ-Aktivierungs-Bedingung den gegenteiligen Effekt zu haben scheint.

9.3 Hypothesenbezogene Ergebnisse

Im Folgenden werden die Überprüfungen der Hypothesen (siehe Kapitel „Zielsetzungen, Fragestellungen und Hypothesen“) zur Selbstwirksamkeitserwartung (SWE), den Leistungen in den EMS-ähnlichen Untertests sowie unter Berücksichtigung der SWE als Kovariate dargelegt.

9.3.1 Selbstwirksamkeitserwartung

Mittels Produkt-Moment-Korrelation nach Pearson wird geprüft, ob es einen signifikanten Zusammenhang zwischen der Selbstwirksamkeitserwartung und den Leistungen in EMS-ähnlichen Untertests gibt $H_1(1)$.

Anhand des Korrelationskoeffizient nach Pearson (r) wird die Richtung und die Enge des Zusammenhangs zwischen zwei Variablen aufgezeigt: Werte von 0.1 bis 0.29 weisen auf einen geringen, Werte von 0.30 bis 0.49 auf einen mittleren und Werte ≥ 0.50 auf einen hohen Zusammenhang hin (Bortz & Döring, 2005).

Tabelle 61 stellt die Zusammenhänge zwischen der Selbstwirksamkeitserwartung und den Leistungen in den sechs EMS-ähnlichen Subtests mittels Produkt-Moment-Korrelation dar.

Tabelle 61: Produkt-Moment-Korrelation, SWE mit den EMS-ähnlichen Untertests (N = 1071)

		EMS-ähnliche Untertests					
		Quantitative u. formale Probleme	Schlauchfiguren	Medizinisch-naturw. Grundv.	Diagramme und Tabellen	Figuren lernen	Fakten lernen
SWE	r	.041	.100**	.038	.067*	.034	.021
	Sig. (2-seitig)	.184	.001	.212	.027	.260	.496

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Aus Tabelle 61 ist zu entnehmen, dass ein geringer positiver, signifikanter Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Leistung im EMS-ähnlichen Untertest *Schlauchfiguren* ($r = .100^{**}$, $p = .001$) und ein sehr geringer positiver, signifikanter Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Leistung im EMS-ähnlichen Subtest *Diagramme und Tabellen* ($r = .067^*$, $p = .027$) besteht.

Die Leistungen in den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Medizinisch naturwissenschaftliches Grundverständnis*, *Figuren lernen* und *Fakten lernen* können als unabhängig von der Selbstwirksamkeitserwartung angenommen werden.

Bezugnehmend auf die Geschlechtergruppen kann nur bei den weiblichen Testpersonen ein sehr geringer positiver, signifikanter Zusammenhang zwischen der Selbstwirksamkeitserwartung und der Leistung im EMS-ähnlichen Subtest *Schlauchfiguren* ($r = .084^*$, $p = .029$) beobachtet werden (siehe Tabelle 62).

Tabelle 62: Produkt-Moment-Korrelation, SWE mit den EMS-ähnlichen Untertests in Abhängigkeit vom Geschlecht (Frauen: N=673, Männer: N = 398)

			EMS-ähnliche Untertests					
			Quantitative u. formale Probleme	Schlauchfiguren	Medizinisch-naturw. Grundv.	Diagramme und Tabellen	Figuren lernen	Fakten lernen
SWE	Frauen	r	.014	.084*	.035	.021	.065	.045
		sig	.713	.029	.366	.593	.092	.244
	Männer	r	.016	.035	.004	.096	-.001	.039
		sig	.743	.487	.937	.056	.977	.432

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Die Prüfung der Hypothese $H_1(2)$ erfolgt mittels dreifaktorieller Varianzanalyse (ANOVA) mit der Selbstwirksamkeitserwartung (SWE-Meanscore) als abhängige Variable und den Faktoren: Geschlecht, Stereotype Threat und Risikoverhalten.

I.) Geschlecht

Die Berechnung der entsprechenden Prüfgröße unter Annahme der Homogenität der Varianzen ($p = .376$) und der Normalverteilung der abhängigen Variable innerhalb jeder Gruppe, ergibt mit $F(1,1047) = 22.354$, $p \leq .0001$ im Faktor Geschlecht ein signifikantes Ergebnis (siehe Tabelle 63).

Tabelle 63: Tests der Zwischensubjekteffekte, AV: Selbstwirksamkeitserwartung

Abhängige Variable: Selbstwirksamkeitserwartung					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Geschlecht	3.61	1	3.61	22.35	≤ .0001
Stereotype Threat	.56	3	.19	1.15	.327
Risikoverhalten	.44	2	.22	1.37	.254
Geschlecht * Stereotype Threat	.90	3	.30	1.86	.135
Geschlecht * Risikoverhalten	.69	2	.34	2.12	.120
Stereotype Threat * Risikoverhalten	1.45	6	.24	1.49	.178
Geschlecht * Stereotype Threat * Risikoverhalten	1.43	6	.24	1.48	.183

Es kann ein signifikanter Unterschied in der Selbstwirksamkeitserwartung (SWE) in Abhängigkeit vom Geschlecht angenommen werden. Anhand von Tabelle 64 wird ersichtlich, dass die männlichen Untersuchungsteilnehmer einen höheren SWE-Meanscore aufweisen als die weiblichen Testpersonen. Die gesamte Tabelle zur Deskriptivstatistik ist im Anhang ersichtlich.

Tabelle 64: Deskriptivstatistik SWE-Meanscore in Abhängigkeit vom Geschlecht

	Geschlecht	N	M	SD	Min	Max
SWE - Meanscore	männlich	398	3.01	0.41	1.40	4.00
	weiblich	673	2.88	0.40	1.60	3.90

Die durchgeführte Effektstärkenberechnung ergibt mit $d = 0.32$ den Hinweis auf einen kleinen Unterschied unter Berücksichtigung der Effektgrößenklassifikation (vgl. Bortz & Döring, 2005, S. 604). Es kann somit in der Selbstwirksamkeitserwartung ein kleiner Effekt zugunsten der männlichen Untersuchungsteilnehmer angenommen werden.

II.) Stereotype Threat

Die Hypothesenprüfung zum Faktor Stereotype Threat ergibt mit $F(3,1047) = 1.153$, $p = .327$ ein nicht signifikantes Ergebnis (siehe Tabelle 63). Es kann kein signifikanter Unterschied in der Selbstwirksamkeitserwartung in Abhängigkeit von den vier Stereotype Threat-Bedingungen (Aktivierung, neutrale Bedingung, Aufklärung und Negativ-Aktivierung) angenommen werden.

III.) Risikoverhalten

Die Berechnung der entsprechenden Prüfgröße zum Faktor Risikoverhalten zeigt mit $F(2,1047) = 1.371$, $p = .254$ (siehe Tabelle 63), dass kein signifikanter Unterschied in der Selbstwirksamkeitserwartung zwischen den drei Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) beobachtet werden kann.

IV.) Wechselwirkungen

Die Hypothesenprüfung ergibt, dass in der Selbstwirksamkeitserwartung keine signifikante Wechselwirkung zwischen *Geschlecht * Stereotype Threat* $F(3,1047) = 1.861$, $p = .135$, *Geschlecht * Risikoverhalten* $F(2,1047) = 2.121$, $p = .120$ und *Stereotype Threat * Risikoverhalten* $F(6,1047) = 1.492$, $p = .178$ angenommen werden kann (siehe Tabelle 63).

Die Berechnung der entsprechenden Prüfgröße zeigt mit $F(6,1047) = 1.476$, $p = .183$ keine signifikante Interaktion höherer Ordnung aus *Geschlecht x Stereotype Threat x Risikoverhalten* in der Selbstwirksamkeitserwartung auf (siehe Tabelle 63).

Zusammenfassend kann an dieser Stelle angeführt werden, dass ein signifikanter Unterschied in der Selbstwirksamkeitserwartung nur in Abhängigkeit vom Geschlecht angenommen werden kann $H_1(2)$. Männliche Untersuchungsteilnehmer gaben eine höhere Selbstwirksamkeitserwartung an als die weiblichen Testpersonen.

Zur Überprüfung der Hypothese $H_1(3)$ wird ein t-Test für unabhängige Stichproben durchgeführt. Es wird untersucht, ob es einen signifikanten Unterschied in der Selbstwirksamkeitserwartung zwischen der Intention, Medizin zu studieren (ja/nein), gibt.

Tabelle 65: t-Test bei unabhängigen Stichproben, SWE und Intention, Medizin zu studieren

		Levene-Test		T-Test für die Mittelwertgleichheit						
		F	Sig.	T	df	Sig. (2-seitig)	Mittlere Differenz	Standard -fehler der Diff.	95% KI der Differenz	
									Untere	Obere
SWE	Varianzen sind gleich	.226	.634	-2.996	1045	.003	-.09832	.03281	-.16271	-.03393
	Varianzen sind nicht gleich			-2.983	265.617	.003	-.09832	.03296	-.16321	-.03343

Die Berechnung der entsprechenden Prüfgröße, unter Annahme der Homogenität der Varianzen ($p = .634$), ergibt mit $t(1045) = -2.996$, $p = .003$ ein signifikantes Ergebnis (siehe Tabelle 65). Es kann ein signifikanter Unterschied in der Selbstwirksamkeitserwartung in Abhängigkeit von der Intention, ein Medizinstudium zu beginnen, beobachtet werden.

Unter Berücksichtigung der Mittelwerte (siehe Tabelle 66) lässt sich erkennen, dass die UntersuchungsteilnehmerInnen, die ein Medizinstudium beginnen möchten einen höheren Wert im SWE-Meanscore aufweisen.

Tabelle 66: Gruppenstatistik SWE in Abhängigkeit von der Intention, Medizin zu studieren

	Intention, Medizin zu studieren	N	M	SD
SWE – Meanscore	nein	863	2.91	0.40
	ja	184	3.01	0.41

Die Berechnung der standardisierten Effektgröße weist mit $d = 0.25$ auf einen geringen Unterschied hin (vgl. Bortz & Döring, 2005, S. 604). Demnach kann ein kleiner Effekt, dass die SchülerInnen, welche ein Medizinstudium beginnen möchten, eine höhere Selbstwirksamkeitserwartung angeben, angenommen werden.

9.3.2 Leistungen in EMS-ähnlichen Untertests

Die Prüfung der Hypothesen $H_1(4-10)$ erfolgt mittels dreifaktorieller multivariater Varianzanalyse (MANOVA) mit den sechs EMS-ähnlichen Untertests als abhängige Variablen und den Zwischensubjektfaktoren: Geschlecht, Stereotype Threat und Risikoverhalten.

Das Ergebnis der Levene-Tests fällt mit $p \geq .133$ jeweils nicht signifikant aus, somit kann von der Homogenität der Varianzen in den sechs EMS-ähnlichen Subtests ausgegangen werden (siehe Tabelle 67).

Die multivariate Normalverteilung der abhängigen Variablen innerhalb der Gruppen kann angenommen werden und die Überprüfung der Homogenität der Varianz-Kovarianz Matrizen mittels Box's-M Test ergibt mit $p = .162$ ein nicht signifikantes Ergebnis, womit auch diese Voraussetzung zur Berechnung der MANOVA gegeben ist.

Tabelle 67: Levene´s Test, EMS-ähnliche Untertests

Levene-Test auf Gleichheit der Fehlervarianzen		
EMS-ähnliche Untertests	F(23,1069)	Sig.
Quantitative und formale Probleme	1.15	.284
Schlauchfiguren	1.33	.140
Med.-naturw. Grundverständnis	1.24	.202
Diagramme und Tabellen	1.33	.137
Figuren lernen	1.01	.454
Fakten lernen	1.34	.133

I.) Geschlecht

Anhand von Tabelle 68 werden die Ergebnisse der multivariaten Varianzanalyse in Abhängigkeit vom Geschlecht dargestellt. Es ist zu entnehmen, dass in fünf EMS-ähnlichen Untertests signifikante Leistungsunterschiede in Abhängigkeit vom Geschlecht beobachtet werden.

Tabelle 68: Tests der Zwischensubjekteffekte, Quelle: Geschlecht

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(1,1069)	Sig.
Quant. und formale Probleme	58.18	58.18	21.49	≤ .0001
Schlauchfiguren	594.12	594.12	53.44	≤ .0001
Med.-naturw. Grundverständnis	19.98	19.98	7.12	.008
Diagramme und Tabellen	27.13	27.13	10.48	.001
Figuren lernen	3.33	3.33	0.80	.373
Fakten lernen	62.15	62.15	20.22	≤ .0001

Die Hypothese $H_1(4)$ kann in den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis*, *Diagramme und Tabellen* sowie *Fakten lernen* bestätigt werden.

Tabelle 69 zeigt auf, dass männliche Untersuchungsteilnehmer in den EMS-ähnlichen Subtests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis* und *Diagramme und Tabellen* höhere Leistungen als ihre weiblichen Kolleginnen erzielen. Im EMS-ähnlichen Subtest *Fakten*

lernen erbringen weibliche Untersuchungsteilnehmer eine höhere Leistung als die männlichen Testpersonen und im EMS-ähnlichen Untertest *Figuren lernen* sind keine Leistungsunterschiede anzunehmen.

Tabelle 69: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit vom Geschlecht (Frauen: N = 684, Männer: N = 409)

			<i>M</i>	<i>SD</i>	Min.	Max.
Geschlecht	Frauen	Quantitative u. formale	2.80	1.57	0	8
		Schlauchfiguren	9.58	3.37	0	20
		Med.-naturw. Grundverständnis	2.69	1.68	0	9
		Diagramme und Tabellen	2.87	1.60	0	9
		Figuren lernen	4.47	2.03	0	8
		Fakten lernen	4.42	1.75	0	8
	Männer	Quantitative u. formale	3.34	1.78	0	10
		Schlauchfiguren	11.21	3.34	0	20
		Med.-naturw. Grundverständnis	3.02	1.69	0	8
		Diagramme und Tabellen	3.24	1.67	0	9
		Figuren lernen	4.36	2.09	0	8
		Fakten lernen	3.94	1.76	0	8

II.) Stereotype Threat

Die Hypothesenprüfung $H_1(5)$ ergibt, dass keine signifikanten Unterschiede ($p \geq .072$) in den Leistungen der EMS-ähnlichen Untertests zwischen den vier Stereotype Threat-Bedingungen (Aktivierung, Neutral, Aufklärung, Negativ-Aktivierung) beobachtet werden, wie mittels Tabelle 70 ersichtlich ist.

Tabelle 70: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	$F(3,1069)$	Sig.
Quant. und formale Probleme	8.69	2.90	1.07	.361
Schlauchfiguren	41.91	13.97	1.26	.288
Med.-naturw. Grundverständnis	8.72	2.91	1.04	.376
Diagramme und Tabellen	8.51	2.84	1.10	.350
Figuren lernen	9.64	3.21	0.77	.512
Fakten lernen	21.52	7.18	2.34	.072

III.) Risikoverhalten

Die Hypothesenprüfung $H_1(6)$ zum Faktor Risikoverhalten zeigt, dass die EMS-ähnlichen Untertests *Quantitative und formale Probleme* ($p = .001$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .006$) sowie *Diagramme und Tabellen* ($p = .003$) in Abhängigkeit von den Risikoverhalten-Bedingungen (Raten, ohne Instruktion, sicheres Antworten) signifikant unterschiedlich bearbeitet werden (siehe Tabelle 71).

Tabelle 71: Tests der Zwischensubjekteffekte, Quelle: Risikoverhalten

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(2,1069)	Sig.
Quant. und formale Probleme	36.90	18.45	6.81	.001
Schlauchfiguren	4.58	2.29	0.21	.814
Med.-naturw. Grundverständnis	28.66	14.33	5.10	.006
Diagramme und Tabellen	29.62	14.81	5.72	.003
Figuren lernen	3.66	1.83	0.44	.646
Fakten lernen	11.49	5.74	1.87	.155

Zur Überprüfung der Unterschiede zwischen den Gruppen werden a-priori Kontraste herangezogen, da angenommen wird, dass die „Rate-Bedingung“ (versuche zu raten, wenn du die richtige Lösung nicht weißt) (vgl. Ben-Shakhar & Sinai, 1991) die Leistungen in den EMS-ähnlichen Untertests begünstigt.

Die Hypothesenprüfung zeigt, dass sich die „Rate-Bedingung“ (Gruppe 1) signifikant von der Bedingung „ohne Instruktion“ (Gruppe 2) bei den EMS-ähnlichen Untertests *Quantitative und formale Probleme* ($p = .001$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .002$) und *Diagramme und Tabellen* ($p = .006$), unterscheidet. Darüber hinaus lässt sich ein signifikanter Unterschied zwischen der „Rate-Bedingung“ (Gruppe 1) und der Bedingung „sicheres Antworten“ (Gruppe 3) bei den EMS-ähnlichen Untertests *Quantitative und formale Probleme* ($p = .002$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .029$) und *Diagramme und Tabellen* ($p = .003$) feststellen.

Aus Tabelle 72 kann entnommen werden, dass bei der Rate-Bedingung in den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Medizinisch-naturwissenschaftliches Grundverständnis* und *Diagramme und Tabellen* höhere Mittelwerte im Gegensatz zu den beiden anderen Risikoverhalten-Bedingungen auftreten.

Tabelle 72: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit von den Risikoverhalten-Bedingungen

Abhängige Variable	Risikoverhalten	<i>M</i>	<i>SD</i>	Min.	Max.
Quantitative und formale Probleme	raten	3.24	1.68	0	10
	ohne Instruktion	2.90	1.68	0	9
	sicher	2.86	1.65	0	8
Schlauchfiguren	raten	10.09	3.60	0	20
	ohne Instruktion	10.31	3.41	2	19
	sicher	10.18	3.32	0	20
Medizinisch-naturwissenschaftliches Grundverständnis	raten	3.02	1.75	0	8
	ohne Instruktion	2.65	1.64	0	7
	sicher	2.77	1.66	0	9
Diagramme und Tabellen	raten	3.21	1.66	0	9
	ohne Instruktion	2.95	1.55	0	9
	sicher	2.85	1.67	0	7
Figuren lernen	raten	4.45	1.97	0	8
	ohne Instruktion	4.54	2.09	0	8
	sicher	4.30	2.10	0	8
Fakten lernen	raten	4.31	1.65	0	8
	ohne Instruktion	4.30	1.81	0	8
	sicher	4.11	1.83	0	8

IV.) Wechselwirkungen

Die Überprüfung der Hypothese $H_1(7)$ ergibt, dass in den Leistungen der EMS-ähnlichen Untertests keine signifikante Wechselwirkung zwischen Geschlecht und Stereotype Threat ($p \geq .673$) angenommen werden kann, wie Tabelle 73 verdeutlicht.

Tabelle 73: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	<i>F</i> (3,1069)	Sig.
Quant. und formale Probleme	4.17	1.39	0.51	.673
Schlauchfiguren	0.43	0.14	0.01	.998
Med.-naturw. Grundverständnis	2.46	0.82	0.29	.831
Diagramme und Tabellen	2.66	0.89	0.34	.794
Figuren lernen	3.30	1.10	0.26	.852
Fakten lernen	1.64	0.55	0.18	.912

Die Hypothesenprüfung $H_1(8)$ zeigt, dass in den beiden EMS-ähnlichen Untertests *Schlauchfiguren* ($p = .004$) und *Figuren lernen* ($p = .003$) signifikante Wechselwirkungen zwischen Geschlecht und Risikoverhalten beobachtet werden (siehe Tabelle 74).

Tabelle 74: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Risikoverhalten

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(2,1069)	Sig.
Quant. und formale Probleme	4.54	2.27	0.84	.432
Schlauchfiguren	125.29	62.64	5.63	.004
Med.-naturw. Grundverständnis	11.11	5.56	1.98	.139
Diagramme und Tabellen	7.53	3.76	1.45	.234
Figuren lernen	48.09	24.05	5.74	.003
Fakten lernen	7.68	3.84	1.25	.287

Anhand der Abbildung 27 sowie Abbildung 28 werden die signifikanten Wechselwirkungen in den Leistungen der EMS-ähnlichen Untertests *Schlauchfiguren* und *Figuren lernen* in Abhängigkeit vom Geschlecht und den Risikoverhalten-Bedingungen dargestellt.

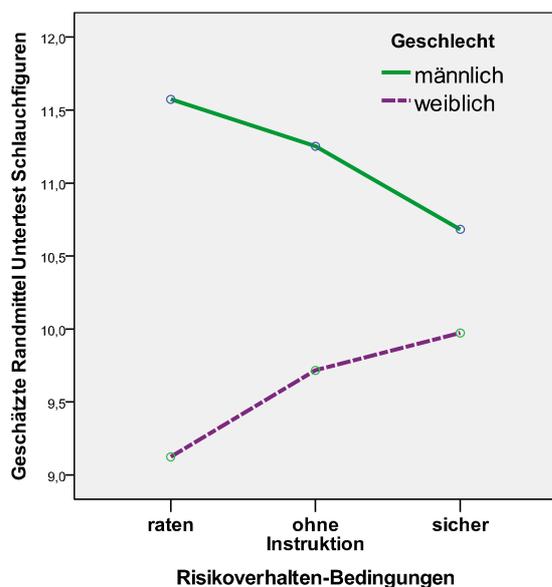


Abbildung 27: Wechselwirkung Geschlecht * Risikoverhalten im UT Schlauchfiguren

Die Abbildung 27 weist darauf hin, dass sich im EMS-ähnlichen Untertest *Schlauchfiguren* die „Rate-Bedingung“ bei den männlichen Untersuchungsteilnehmern positiv auf die Leistung auswirkt, wohingegen diese Instruktion bei den weiblichen Testpersonen den gegenteiligen Effekt zu haben scheint.

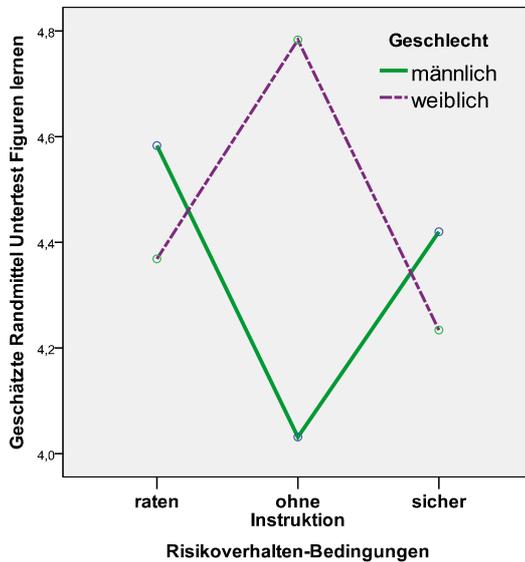


Abbildung 28: Wechselwirkung Geschlecht * Risikoverhalten im UT Figuren lernen

Anhand von Abbildung 28 ist erkennbar, dass im EMS-ähnlichen Untertest *Figuren lernen* die Risikoverhalten-Bedingung „ohne Instruktion“ bei den weiblichen Schülerinnen einen positiven Effekt auf die Leistungen auszuwirken scheint, während bei den männlichen Testpersonen genau das Gegenteil beobachtet werden kann.

Die Überprüfung der $H_1(9)$ zeigt, dass im EMS-ähnlichen Subtest *Diagramme und Tabellen* ($p = .021$) eine signifikante Wechselwirkung zwischen Stereotype Threat und Risikoverhalten angenommen werden kann, wie Tabelle 75 zu entnehmen ist.

Tabelle 75: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat * Risikoverhalten

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(6,1069)	Sig.
Quant. und formale Probleme	26.36	4.39	1.62	.137
Schlauchfiguren	79.22	13.20	1.19	.310
Med.-naturw. Grundverständnis	27.14	4.52	1.61	.141
Diagramme und Tabellen	38.73	6.46	2.49	.021
Figuren lernen	16.93	2.82	0.67	.671
Fakten lernen	5.09	0.85	0.28	.948

Die Abbildung 29 zeigt die signifikante Wechselwirkung auf und deutet darauf hin, dass im EMS-ähnlichen Untertest *Diagramme und Tabellen* die Faktorstufenkombination „Rate-Bedingung“ mit „Neutral“ den Leistungen der Testpersonen entgegenkommt.

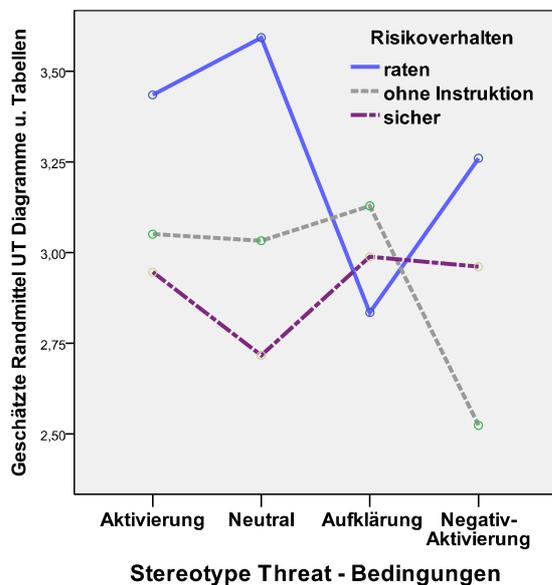


Abbildung 29: Wechselwirkung Stereotype Threat * Risikoverhalten im UT Diagramme und Tabellen

Die Hypothesenprüfung $H_1(10)$ ergibt, dass eine signifikante Interaktion höherer Ordnung aus Geschlecht x Stereotype Threat x Risikoverhalten im EMS-ähnlichen Untertest *Schlauchfiguren* ($p = .027$) beobachtet werden kann (siehe Tabelle 76).

Tabelle 76: Tests der Zwischensubjekteffekte, Quelle: Geschlecht* Stereotype Threat * Risikoverhalten

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(6,1069)	Sig.
Quant. und formale Probleme	23.17	3.86	1.43	.201
Schlauchfiguren	159.32	26.55	2.39	.027
Med.-naturw. Grundverständnis	5.17	0.86	0.31	.934
Diagramme und Tabellen	16.67	2.78	1.07	.377
Figuren lernen	24.10	4.02	0.96	.452
Fakten lernen	17.16	2.86	0.93	.472

Die signifikante Interaktion höherer Ordnung im EMS-ähnlichen Untertest Schlauchfiguren wird nachfolgend dargestellt (siehe Abbildung 30 und Abbildung 31).

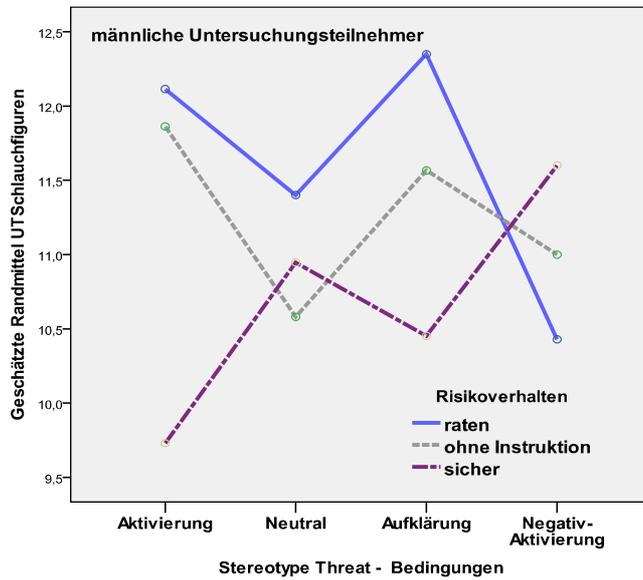


Abbildung 30: Interaktion höherer Ordnung im UT Schlauchfiguren, männliche Testpersonen

Die Abbildung 30 zeigt, dass im EMS-ähnlichen Untertest Schlauchfiguren die Faktorstufenkombination „Rate-Bedingung“ mit „Aufklärung“ bei männlichen Untersuchungsteilnehmern positiv auf die Leistung wirkt.

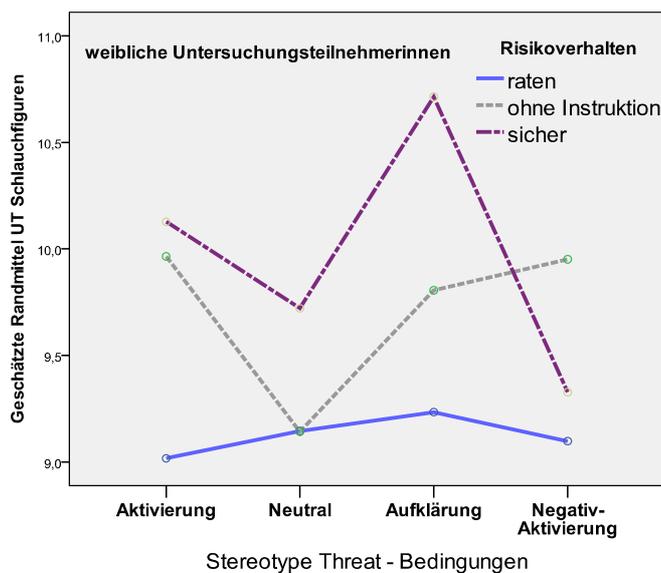


Abbildung 31: Interaktion höherer Ordnung im UT Schlauchfiguren, weibliche Testpersonen

Anhand Abbildung 31 ist erkennbar, dass im EMS-ähnlichen Untertest *Schlauchfiguren* die Faktorstufenkombination „sicheres Antworten“ mit „Aufklärung“ bei den weiblichen Untersuchungsteilnehmern einen positiven Effekt auf die Leistung auszuüben scheint.

9.3.3 Leistungen in EMS-ähnlichen Untertests unter Berücksichtigung der SWE als Kovariate

In diesem Abschnitt werden die Hypothesenprüfungen $H_1(11-17)$ mittels der dreifaktoriellen multivariaten Kovarianzanalyse (MANCOVA) dargestellt.

Die MANCOVA wird mit der Selbstwirksamkeitserwartung (SWE) als Kovariate, den Leistungen in den EMS-ähnlichen Untertests als abhängige Variablen und den Zwischensubjektfaktoren Geschlecht, Stereotype Threat und Risikoverhalten durchgeführt.

Auf Grund der nicht signifikanten Ergebnisse der Levene-Tests ($p \geq .078$) ist die Voraussetzung der Homogenität der Varianzen in den sechs EMS-ähnlichen Subtests gegeben (siehe Tabelle 77). Die Homogenität der Regressionsgeraden kann angenommen werden, der Box's-M Test zur Prüfung der Homogenität der Varianz-Kovarianz-Matrizen ergibt mit $p = .206$ ein nicht signifikantes Ergebnis.

Tabelle 77: Levene's Test, EMS-ähnliche Untertests (unter Berücksichtigung SWE)

Levene-Test auf Gleichheit der Fehlervarianzen		
EMS-ähnliche Untertests	F(23, 1047)	Sig.
Quantitative und formale Probleme	1.26	.181
Schlauchfiguren	1.30	.155
Med.-naturwissenschaftliches Grundverständnis	1.15	.283
Diagramme und Tabellen	1.45	.078
Figuren lernen	0.76	.780
Fakten lernen	1.31	.147

Anhand von Tabelle 78 ist zu entnehmen, dass ein signifikanter Einfluss (auf Basis der MANCOVA) der Selbstwirksamkeitserwartung auf die Leistungen im EMS-ähnlichen Untertest *Schlauchfiguren* ($p = .015$) angenommen werden kann.

Tabelle 78: Tests der Zwischensubjekteffekte, Quelle: Selbstwirksamkeitserwartung

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(1,1046)	Sig.
Quant. und formale Probleme	0.18	0.18	0.07	.797
Schlauchfiguren	65.38	65.38	5.94	.015
Med.-naturw. Grundverständnis	1.53	1.53	0.55	.460
Diagramme und Tabellen	3.19	3.19	1.25	.264
Figuren lernen	7.55	7.55	1.84	.176
Fakten lernen	7.64	7.64	2.59	.108

I.) Geschlecht

Mittels Tabelle 79 werden die Ergebnisse der multivariaten Kovarianzanalyse in Abhängigkeit vom Geschlecht dargestellt. Es zeigt sich, dass (analog zur MANOVA) in den Leistungen von fünf EMS-ähnlichen Subtests signifikante Unterschiede zwischen den männlichen und weiblichen Testpersonen unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate angenommen werden können.

Tabelle 79: Tests der Zwischensubjekteffekte, Quelle: Geschlecht, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(1,1046)	Sig.
Quant. und formale Probleme	57.86	57.86	21.47	≤ .0001
Schlauchfiguren	520.53	520.53	47.33	≤ .0001
Med.-naturw. Grundverständnis	17.80	17.80	6.37	.012
Diagramme und Tabellen	26.35	26.35	10.34	.001
Figuren lernen	5.73	5.73	1.39	.238
Fakten lernen	67.14	67.14	22.73	≤ .0001

Demgemäß ergibt die Hypothesenprüfung $H_1(11)$ zum Faktor Geschlecht unter Berücksichtigung der SWE als Kovariate, dass die EMS-ähnlichen Untertests *Quantitative und formale Probleme* ($p \leq .0001$), *Schlauchfiguren* ($p \leq .0001$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .012$), *Diagramme und Tabellen* ($p = .001$) sowie *Fakten lernen* ($p \leq .0001$) signifikant unterschiedlich bearbeitet werden.

Unter Berücksichtigung der Mittelwerte (siehe Tabelle 80) zeigt sich, dass männliche Testpersonen in den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis* und *Diagramme und Tabellen* höhere Leistungen erbringen als die weiblichen Untersuchungsteilnehmer. Im Subtest *Fakten lernen* erzielen die Schülerinnen eine höhere Leistung als ihre männlichen Kollegen.

Tabelle 80: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit vom Geschlecht unter Berücksichtigung SWE (Frauen: N = 673, Männer: N = 398)

		<i>M</i>	<i>SD</i>	
Geschlecht	Frauen	Quantitative u. formale	2.82	1.57
		Schlauchfiguren	9.60	3.35
		Med.-naturw. Grundverständnis	2.71	1.67
		Diagramme und Tabellen	2.89	1.59
		Figuren lernen	4.49	2.02
		Fakten lernen	4.45	1.71
	Männer	Quantitative u. formale	3.38	1.77
		Schlauchfiguren	11.25	3.34
		Med.-naturw. Grundverständnis	3.05	1.69
		Diagramme und Tabellen	3.29	1.65
		Figuren lernen	4.36	2.07
		Fakten lernen	3.97	1.73

II.) Stereotype Threat

Unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate zeigt die Hypothesenprüfung $H_1(12)$, dass ein signifikanter Unterschied in den Leistungen des EMS-ähnlichen Untertests *Fakten lernen* ($p = .048$) zwischen den Stereotype Threat-Bedingungen (Aktivierung, Neutral, Aufklärung, Negativ-Aktivierung) angenommen werden kann (siehe Tabelle 81).

Bei der MANOVA wurden hier noch keine signifikanten Unterschiede beobachtet.

Tabelle 81: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(3,1046)	Sig.
Quant. und formale Probleme	8.75	2.92	1.08	.356
Schlauchfiguren	49.39	16.46	1.50	.214
Med.-naturw. Grundverständnis	7.89	2.63	0.94	.420
Diagramme und Tabellen	6.44	2.15	0.84	.471
Figuren lernen	8.23	2.74	0.67	.572
Fakten lernen	23.48	7.83	2.65	.048

Die Prüfung auf Unterschiede zwischen den einzelnen Stereotype Threat-Bedingungen bei dem EMS-ähnlichen Untertest *Fakten lernen* wird mittels a priori Kontrasten durchgeführt. Die Hypothesenprüfung zeigt, dass nur der Vergleich der Bedingungen „Aktivierung“ und „Neutral“ bei dem EMS-ähnlichen Subtest *Fakten lernen* ein signifikantes Ergebnis ($p = .022$) erbringt. Die Mittelwerte weisen darauf hin, dass in der Bedingung „Aktivierung“ ($M = 4.33$, $SD = 1.78$) höhere Leistungen im EMS-ähnlichen Untertest *Fakten lernen* gegenüber der Bedingung „Neutral“ ($M = 4.06$, $SD = 1.80$) beobachtet werden können, sobald der Einfluss der Kovariate Selbstwirksamkeitserwartung herausgerechnet (vgl. Bortz & Döring, 2005) wird.

III.) Risikoverhalten

Die Hypothesenprüfung $H_1(13)$ zum Faktor Risikoverhalten unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate ergibt (analog zur MANOVA) signifikante Unterschiede in den EMS-ähnlichen Untertests *Quantitative und formale Probleme* ($p = .001$), *Medizinisch-naturwissenschaftliches Grundverständnis* ($p = .006$) und *Diagramme und Tabellen* ($p = .004$), wie Tabelle 82 zu entnehmen ist.

Tabelle 82: Tests der Zwischensubjekteffekte, Quelle: Risikoverhalten, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(2,1046)	Sig.
Quant. und formale Probleme	39.46	19.73	7.32	.001
Schlauchfiguren	7.18	3.59	0.33	.722
Med.-naturw. Grundverständnis	28.88	14.44	5.17	.006
Diagramme und Tabellen	28.48	14.24	5.59	.004
Figuren lernen	2.64	1.32	0.32	.726
Fakten lernen	9.79	4.89	1.66	.191

IV.) Wechselwirkungen

Die Hypothesenprüfung $H_1(14)$ ergibt, dass unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate (analog zur MANOVA) keine signifikante Wechselwirkung zwischen Geschlecht und Stereotype Threat ($p \geq .695$) in den Leistungen der EMS-ähnlichen Untertests beobachtet werden kann, wie Tabelle 83 zeigt.

Tabelle 83: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(3,1046)	Sig.
Quant. und formale Probleme	3.75	1.25	0.46	.707
Schlauchfiguren	0.46	0.15	0.01	.998
Med.-naturw. Grundverständnis	4.04	1.35	0.48	.695
Diagramme und Tabellen	1.21	0.40	0.16	.924
Figuren lernen	3.82	1.27	0.31	.819
Fakten lernen	2.26	0.75	0.26	.858

Die Überprüfung der Hypothese $H_1(15)$ unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate ergibt, dass (analog zur MANOVA) signifikante Wechselwirkungen zwischen Geschlecht und Risikoverhalten in den EMS-ähnlichen Untertests *Schlauchfiguren* ($p = .003$) und *Figuren lernen* ($p = .001$) angenommen werden können (siehe Tabelle 84).

Tabelle 84: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Risikoverhalten, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(2,1046)	Sig.
Quant. und formale Probleme	445	2.23	0.83	.438
Schlauchfiguren	125.68	62.84	5.71	.003
Med.-naturw. Grundverständnis	10.31	5.16	1.85	.159
Diagramme und Tabellen	7.73	3.87	1.52	.220
Figuren lernen	57.65	28.83	7.01	.001
Fakten lernen	7.06	3.53	1.19	.303

Unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate zeigt die Hypothesenprüfung $H_1(16)$ im EMS-ähnlichen Untertest *Diagramme und Tabellen* ($p = .011$) eine signifikante Wechselwirkung zwischen Stereotype Threat und Risikoverhalten (analog zur MANOVA) auf, wie Tabelle 85 zu entnehmen ist.

Tabelle 85: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat * Risikoverhalten, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(6,1046)	Sig.
Quant. und formale Probleme	22.24	3.71	1.38	.221
Schlauchfiguren	83.19	13.87	1.26	.273
Med.-naturw. Grundverständnis	29.03	4.84	1.73	.111
Diagramme und Tabellen	42.80	7.13	2.80	.011
Figuren lernen	15.57	2.60	0.63	.705
Fakten lernen	2.01	0.33	0.11	.995

Die Hypothesenprüfung $H_1(17)$ ergibt unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate, dass nunmehr keine signifikante Interaktion höherer Ordnung in den Leistungen der EMS-ähnlichen Untertests aus Geschlecht x Stereotype Threat x Risikoverhalten ($p \geq .064$) angenommen werden kann (siehe Tabelle 86).

Bei der MANOVA konnte im EMS-ähnlichen Untertest *Schlauchfiguren* eine signifikante Interaktion höherer Ordnung aus Geschlecht x Stereotype Threat x Risikoverhalten beobachtet werden.

Tabelle 86: Tests der Zwischensubjekteffekte, Quelle: Geschlecht *Stereotype Threat * Risikoverhalten, Kovariate SWE

EMS-ähnliche Untertests	Quadratsumme vom Typ III	Mittel der Quadrate	F(6,1046)	Sig.
Quant. und formale Probleme	22.40	3.73	1.39	.217
Schlauchfiguren	131.53	21.92	1.99	.064
Med.-naturw. Grundverständnis	5.27	0.88	0.31	.930
Diagramme und Tabellen	12.73	2.12	0.83	.545
Figuren lernen	26.33	4.39	1.07	.380
Fakten lernen	19.81	3.30	1.12	.350

9.3.4 Risikoverhaltens-Tendenz

Die Überprüfung der Wirksamkeit der „Risikoverhalten-Bedingungen“ (Raten, ohne Instruktion und sicheres Antworten) mittels des Game of Dice Task (GDT) (Brand et al., 2005) konnte auf Basis von 471 Personen untersucht werden. Anhand von Tabelle 87 wird die Deskriptivstatistik des GDT-Nettoscores gesamt und hinsichtlich der Risikoverhalten-Bedingungen dargestellt. Hierbei ist darauf hinzuweisen, dass niedrigere Werte im GDT für eine ausgeprägtere Risikoverhaltens-Tendenz stehen.

Tabelle 87: Deskriptivstatistik GDT (Gesamt und Risikoverhalten-Bedingungen)

		Gesamt ns_GDT	raten	ohne Instruktion	sicher
N	Gültig	471	157	163	151
	Fehlend	622	216	205	201
M		4.31	2.72	5.28	4.91
SD		9.71	9.67	9.31	10.01
Schiefe		-.516	-.414	-.605	-.561
Kurtosis		-.645	-.745	-.473	-.636
Min.		-18	-18	-16	-18
Max.		18	18	18	18

Die Normalverteilung der Daten je Gruppe kann anhand der Werte der Schiefe und der Kurtosis angenommen werden, sodass die Prüfung auf Unterschiedlichkeit ($H_1(18)$) mittels einfaktorieller Varianzanalyse (ANOVA) zulässig ist.

Die Berechnung der entsprechenden Prüfgröße unter Annahme der Homogenität der Varianzen ($p = .610$) ergibt mit $F(2,468) = 3.238$, $p = .040$ ein signifikantes Ergebnis. Die Risikoverhaltens-Tendenz (im GDT) fällt in Abhängigkeit von der Risikoverhalten-Bedingung signifikant unterschiedlich aus, somit kann die Hypothese $H_1(18)$ bestätigt werden.

Die Überprüfung der Unterschiede zwischen den drei Gruppen erfolgt mittels a-priori Kontrasten, da angenommen wird, dass die „Rate-Bedingung“ der Risikoverhaltens-Tendenz entgegenkommt. Die Hypothesenprüfung zeigt, dass sich die Rate-Bedingung

(Gruppe 1) signifikant von den beiden anderen Bedingungen (ohne Instruktion, Gruppe 2 und sicheres Antworten, Gruppe 3) unterscheidet (siehe Tabelle 88).

Tabelle 88: Prüfgrößen und Parameter zum Gruppenvergleich mittels a-priori Kontrasten, GDT

	Kontrast (Gruppen)	Kontrastwert	SE	t(468)	Signifikanz (2-seitig)
Nettoscore GDT	1 (1 vs 2)	2.56	1.080	2.367	.018
	2 (1 vs 3)	2.19	1.101	1.993	.047
	3 (2 vs 3)	-.36	1.091	-.332	.740

Abbildung 32 veranschaulicht mittels Boxplot-Diagrammen die Wirksamkeit der „Risikoverhalten-Bedingungen. Die Rate-Bedingung weist eine signifikant ausgeprägtere Risikoverhaltens-Tendenz (NS-GDT) auf. Die beiden anderen Bedingungen unterscheiden sich nicht voneinander.

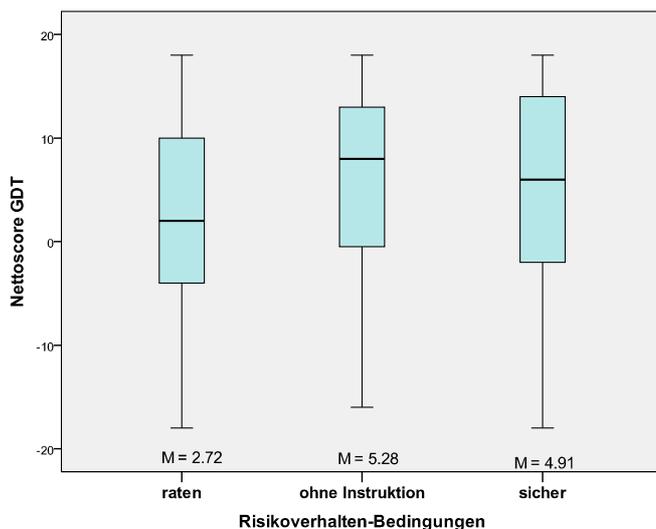


Abbildung 32: Boxplot Diagramme GDT

Darüber hinaus wird mittels t - Test für unabhängige Stichproben untersucht, ob es einen signifikanten Unterschied in der Risikoverhaltens-Tendenz zwischen männlichen und weiblichen Testpersonen gibt (H_1 19).

Anhand von Tabelle 89 wird ersichtlich, dass Männer im GDT einen niedrigeren Mittelwert erreichen als Frauen.

Tabelle 89: Gruppenstatistik NS-GDT

	Geschlecht	N	M	SD
NS-GDT	männlich	183	3.25	9.88
	weiblich	288	4.98	9.55

Die Berechnung der entsprechenden Prüfgröße unter Annahme der Homogenität der Varianzen ($p = .382$) ergibt mit $t(469) = -1.898$, $p = .058$ ein nicht signifikantes Ergebnis (siehe Tabelle 90). Die Hypothese $H_1(19)$ kann somit nicht bestätigt werden.

Tabelle 90: t-Test bei unabhängigen Stichproben, Risikoverhaltens-Tendenz und Geschlecht

		Levene-Test		T-Test für die Mittelwertgleichheit						
		F	Sig.	T	df	Sig. (2-seitig)	Mittlere Differenz	Standar d-fehler der Diff.	95% KI der Differenz	
									Untere	Obere
NS-GDT	Varianzen sind gleich	.766	.382	-1.898	469	.058	-1.737	.915	-3.535	.061
	Varianzen sind nicht gleich			-1.884	377.600	.060	-1.737	.922	-3.550	.076

10. Diskussion

In der vorliegenden Untersuchung wurde der Einfluss von Geschlecht, Stereotype Threat und Risikoverhalten auf die Leistungen bei sechs EMS-ähnlichen Untertests analysiert, wobei gleichzeitig dem Effekt des Konstruktes der Selbstwirksamkeitserwartung ein zentrales Interesse zukam.

Die Ergebnisse verdeutlichen zunächst, dass in fünf EMS-ähnlichen Untertests geschlechtsspezifische Leistungsunterschiede angenommen werden können. Übereinstimmend mit den EMS-Ergebnissen von Mallinger et al. (2010), erzielten die männlichen Testpersonen in den EMS-ähnlichen Subtests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis* und *Diagramme und Tabellen* höhere Leistungen im Vergleich zu den weiblichen Untersuchungsteilnehmern. Geschlechtsspezifische Leistungsunterschiede zugunsten der weiblichen Schüler konnten nur im EMS-ähnlichen Untertest *Fakten lernen* festgestellt werden. Dieser Befund entspricht dem Ergebnis des EMS-Testeinsatzes 2010 in Österreich gemäß Mallinger et al. (2010) mit der Ausnahme, dass die weiblichen Studienbewerber in einem weiteren Untertest, nämlich *Figuren lernen*, höhere Leistungen als die männlichen Studienanwärter erzielten. Insgesamt spiegeln somit die Untersuchungsergebnisse anhand der sechs vorgegebenen EMS-ähnlichen Untertests, die im EMS aufgezeigten geschlechtsspezifischen Leistungsunterschiede von Mallinger et al. (2010) wider, mit Ausnahme des EMS-ähnlichen Untertests *Figuren lernen*, wobei kein Leistungsunterschied zugunsten der weiblichen Schüler beobachtet werden konnte.

Spiel et al. (2008) stellten im Rahmen ihrer Evaluationsstudie des EMS fest, dass StudienbewerberInnen, die den EMS erfolgreich absolvierten ihre mathematisch-naturwissenschaftlichen Fähigkeiten höher einschätzten als BewerberInnen, die nicht zum Medizinstudium zugelassen wurden. In diesem Zusammenhang lässt sich bei Betrachtung der *persönlichen Fähigkeitszuschreibungen* der vorliegenden Untersuchung erkennen, dass sich die männlichen Schüler im Vorfeld eine höhere Fähigkeit in Bezug auf Items zu *Quantitative und formale Probleme* sowie *Medizinisch-naturwissenschaftlichem Grundverständnis* zuschrieben im Vergleich zu ihren weiblichen Kolleginnen. Darüber hinaus schätzen die männlichen Untersuchungsteilnehmer auch ihre Fähigkeiten in Bezug auf Items zu

Schlauchfiguren, Diagramme und Tabellen sowie *Figuren lernen* höher ein. Demgemäß weisen die Ergebnisse darauf hin, dass männliche Testpersonen ihre Fähigkeiten hinsichtlich der Bewältigung der Aufgaben im Vorhinein höher beurteilten und mit Ausnahme von *Figuren lernen*, auch eine höhere Leistung bei den entsprechenden EMS-ähnlichen Untertests erzielten.

Angesichts der Einschätzungen zur Selbstwirksamkeitserwartung zeigen mehrere Studien geschlechtsspezifische Unterschiede zugunsten der Männer auf (Hinz et al., 2006; Pajares & Miller, 1994; Schwarzer et al., 1997, 1999). In Einklang mit diesen Befunden konnte in der Untersuchung festgestellt werden, dass sich die männlichen Untersuchungsteilnehmer eine höhere Selbstwirksamkeitserwartung zuschreiben als die weiblichen Testpersonen ($d = 0.32$). Als eine etwaige Erklärung könnten gemäß Schwarzer et al. (1999) Sozialisationsunterschiede zwischen Frauen und Männer fungieren.

Weiters zeichnete sich in den Untersuchungsergebnissen ein signifikanter Unterschied in der Selbstwirksamkeitserwartung in Abhängigkeit von der Intention, Medizin zu studieren, ab. Es stellte sich heraus, dass UntersuchungsteilnehmerInnen, die ein Medizinstudium intendieren eine höhere Selbstwirksamkeitserwartung angaben als SchülerInnen, die kein Medizinstudium anstreben ($d = 0.25$). Es ist zwar nur von einem kleinen Effekt zu sprechen (vgl. Bortz & Döring, 2005, S. 604), jedoch könnte gemäß der Literatur gefolgert werden, dass SchülerInnen, die Medizin studieren wollen womöglich tendenziell motivierter sind sich neuen, herausfordernden Aufgaben zu stellen (Bandura, 1977, 1986, 1997; Jerusalem, 2002; Jerusalem & Mittag, 1999) im Vergleich zu ihren SchulkollegInnen. Die Selbstwirksamkeitserwartung scheint zudem auch mit dem Verbleib in angestrebten Studienrichtungen im positiven Zusammenhang zu stehen (Lent et al., 1984, 1986). Darüber hinaus könnten sich unter den Testpersonen auch SchülerInnen befinden, die etwaig aufgrund geringer Selbstwirksamkeitserwartungen ein Medizinstudium nicht wagen bzw. es sich nicht zutrauen (Bandura, 1977, 1986, 1997; Jerusalem, 2002). In diesem Zusammenhang wurde bereits von Autoren dargelegt, dass die Selbstwirksamkeitserwartung mit den wahrgenommenen beruflichen Optionen positiv zusammenhängt (Bandura, 1986, 1997) respektive auch als Prädiktor angenommen werden kann (Lent et al., 1986).

Im Hinblick auf die Leistung verdeutlichen Untersuchungen einen positiven Zusammenhang zwischen der Selbstwirksamkeitserwartung und beispielsweise den schulischen Leistungen (Edelstein et al., 1997; Jerusalem & Mittag, 1999; Lent et al.,

1984; Luszczynska et al., 2005; Mittag et al., 2002), der Lernfreude (Edelstein et al., 1997; Jerusalem & Mittag, 1999; Mittag et al., 2002) sowie den akademischen Leistungen (Lent et al., 1984, 1986; Schwarzer et al., 1999). Im Rahmen einer Metaanalyse von Multon et al. (1991) konnte eine Effektgröße von $r = 0.38$ für den Zusammenhang zwischen der Selbstwirksamkeitserwartung und Leistung (in Bezug auf standardisierte Tests, schulbezogene Erhebungsmethoden und fähigkeitsspezifische Aufgaben) aufgezeigt werden. In der vorliegenden Untersuchung ließ sich ein positiver, signifikanter Zusammenhang zwischen der Selbstwirksamkeitserwartung und den Leistungen bei den beiden EMS-ähnlichen Subtests *Schlauchfiguren* ($r = .100$) und *Diagramme und Tabellen* ($r = .067$) feststellen, wobei die beobachteten Zusammenhänge allerdings als relativ schwach zu beurteilen sind (vgl. Bortz & Döring, 2005, S. 604). Demnach weist das Ergebnis darauf hin, dass Testpersonen, die sich eine höhere Selbstwirksamkeitserwartung zuweisen, im Subtest *Schlauchfiguren*, der das räumliche Vorstellungsvermögen erfasst (Hänsgen & Spicher, 2010), eine höhere Leistung erzielen. Die Leistungen in den weiteren vorgegebenen EMS-ähnlichen Subtests *Medizinisch naturwissenschaftliches Grundverständnis*, *Figuren lernen*, *Fakten lernen* und *Quantitative und formale Probleme* hingegen, stehen mit der Selbstwirksamkeitserwartung nicht im Zusammenhang.

Bezugnehmend auf das Risikoverhalten, führten Bolger und Kellaghan (1990) als einen möglichen Grund für die im Rahmen ihrer Studie beobachtete geringere Leistung der weiblichen Schülerinnen bei Multiple-Choice Items, die höhere Rate-Tendenz (Risikoverhalten) der männlichen Schüler an. In diesem Zusammenhang konnten Untersuchungen eine höhere Tendenz von weiblichen Testpersonen, Items auszulassen, im Vergleich zu den männlichen Untersuchungsteilnehmern aufzeigen (z.B.: Hanna, 1986). Zudem verdeutlichten Ben-Shakhar und Sinai (1991), dass der Unterschied auch angenommen werden kann, wenn vor der Testbearbeitung vermittelt wird, dass man raten solle, wenn man sich nicht ganz sicher sei. In der vorliegenden Untersuchung konnten signifikante Leistungsunterschiede in Abhängigkeit vom Geschlecht und Risikoverhalten-Bedingungen (Raten, ohne Instruktion und sicheres Antworten) bei zwei EMS-ähnlichen Untertests angenommen werden. Erstens zeigte sich im EMS-ähnlichen Untertest *Schlauchfiguren*, dass die „Rate-Bedingung“ die Leistung bei den männlichen Schülern positiv zu beeinflussen scheint, wohingegen bei den weiblichen Schülerinnen das Gegenteil beobachtet werden kann. Mittels der Bedingung „sicheres Antworten“ (man solle nur dann eine Antwort ankreuzen, wenn

man sich ganz sicher sei) wiederum, zeichneten sich im EMS-ähnlichen Untertest *Schlauchfiguren* positive Effekte auf die Leistung weiblicher Testpersonen und der konträre Effekt bei den männlichen Untersuchungsteilnehmern, ab. Zweitens ließ sich im EMS-ähnlichen Untertest *Figuren lernen*, anhand dessen die Merkfähigkeit von bestimmten Einzelheiten erfasst wird (Hänsgen & Spicher, 2010), ebenso ein positiver Einfluss der „Rate-Bedingung“ auf die Leistung der männlichen Testpersonen feststellen, während die Bedingung „ohne Instruktion“ (arbeite die Aufgaben aufmerksam und konzentriert durch) einen positiven Effekt auf die Leistung der Frauen aufzeigt. Demgemäß scheint die Instruktion zum „Raten“ die Leistung der weiblichen Testteilnehmer in diesen beiden Untertests nicht zu begünstigen, während dies bei den männlichen Untersuchungsteilnehmern schon der Fall zu sein scheint.

Einen weiteren Gegenstand der Untersuchung stellte das Phänomen Stereotype Threat dar. Untersuchungen belegen, dass Frauen im Vergleich zu Männern eine geringere Mathematikleistung (z.B.: Spencer et al., 1999) sowie visuell-räumliche Leistung (Campbell & Collaer, 2009) erzielen, wenn in der jeweiligen Bedingung negative Stereotype (z.B.: „Männer seien in mathematischen Aufgaben besser als Frauen“) aktiviert sind (z.B.: Smith & White, 2002). Im Gegensatz dazu erzielen Frauen ebenso gute Leistungen wie Männer, wenn die Bedrohung durch negative Stereotype beispielsweise mittels der Instruktion (z.B.: „Im Mathematiktest habe es nie geschlechtsspezifische Leistungsunterschiede gegeben“) reduziert wird (z.B.: Spencer et al., 1999). Zudem konnten mehrere Techniken aufgezeigt werden, die Stereotype Threat-Effekte reduzieren, darunter zählt zum Beispiel die Aufklärung der Frauen über dieses Phänomen (Johns et al., 2005). In der vorliegenden Untersuchung konnte zum einen die Wirksamkeit der vier Stereotype Threat-Bedingungen (*Aktivierung, Neutral, Aufklärung und Negativ-Aktivierung*) anhand der *geschlechtsspezifischen Fähigkeitszuschreibungen* untermauert werden. Zum anderen und im Widerspruch zu den dargelegten Studien (z.B.: Spencer et al., 1999) konnte jedoch kein Unterschied in den Leistungen der EMS-ähnlichen Untertests in Abhängigkeit von Geschlecht und Stereotype Threat-Bedingung festgestellt werden. An dieser Stelle ist somit fraglich, warum hier keine signifikante Wechselwirkung beobachtet werden konnte.

Zum einen muss möglicherweise in Betracht gezogen werden, ob die Stichprobe der SchülerInnen der 12. Schulstufe AHS für BewerberInnen eines Medizinstudiums repräsentativ ist. Denn in der Stichprobe der getesteten SchülerInnen (N= 1068 gültige Fälle) befanden sich nur 17,5%, die nach der Matura ein Medizinstudium intendierten. Demgemäß ist unklar, inwieweit sich die anderen SchülerInnen, die kein

Medizinstudium in Betracht zogen, bei der Bearbeitung der EMS-ähnlichen Aufgaben angestrengt haben, zumal ihnen auch keine Leistungsrückmeldung in Aussicht stand. Insofern wäre es in zukünftigen Untersuchungen wünschenswert, SchülerInnen zu testen, die tatsächlich ein Medizinstudium anstreben. In der vorliegenden Untersuchung wäre dies aufgrund des enormen zeitlichen und organisatorischen Aufwands nicht möglich gewesen.

Ein weitere Einschränkung stellen niedrige Reliabilitäten (vgl. Kline, 1999, zitiert nach Field, 2009, S. 675) dar, die im Rahmen der vorliegenden Studie bei vier EMS-ähnlichen Untertests (*Quantitative und formale Probleme, Medizinisch-naturwissenschaftliches Grundverständnis, Diagramme und Tabellen* sowie *Fakten lernen*) festgestellt wurden. Bereits Spiel et al. (2008) beobachteten in ihrer Analyse von neun EMS Untertests keine hohen Reliabilitäten (≤ 0.75) sowie lagen die Trennschärfen von insgesamt 178 analysierten Aufgaben nur bei 40 Items über 0.30. Dieser Befund steht im Einklang mit den Ergebnissen der vorliegenden Studie, auch hier konnte die überwiegende Mehrheit der EMS-ähnlichen Aufgaben als nicht trennscharf (≥ 0.30) (vgl. Bortz & Döring, 2005) angenommen werden.

Im Hinblick auf die Leistungen bei den EMS-ähnlichen Untertests, zeichnete sich eine signifikante Interaktion höherer Ordnung aus Geschlecht x Stereotype Threat x Risikoverhalten im EMS-ähnlichen Subtest *Schlauchfiguren* ab. Bei näherer Betrachtung stellte sich heraus, dass bei den weiblichen Untersuchungsteilnehmern die Faktorstufenkombination „Aufklärung“ (über Stereotype Threat) (vgl. Johns et al., 2005) mit „sicheres Antworten“ (nur ankreuzen, wenn man ganz sicher ist) die Leistungen im EMS-ähnlichen Untertest *Schlauchfiguren* positiv zu beeinflussen scheint. Im Gegensatz dazu ließ sich bei der Faktorstufenkombination „Aktivierung“ (von Stereotype Threat) (vgl. Smith & White, 2002) mit „Raten“ (versuche zu raten, wenn du die Lösung nicht weißt) (vgl. Ben-Shakhar & Sinai, 1991) ein eher beeinträchtigender Effekt auf die Bearbeitung von Items zum räumlichen Vorstellungsvermögen (Hänsgen & Spicher, 2010) bei den weiblichen Testpersonen erkennen. Angesichts der Ergebnisse bei den männlichen Schülern, zeigte einerseits die Faktorstufenkombination „Aufklärung“ mit „Raten“ einen positiven und andererseits die Faktorstufenkombination „Aktivierung“ mit „sicheres Antworten“ einen negativen Einfluss auf die Leistungen im EMS-ähnlichen Untertest *Schlauchfiguren*. Gemäß den Untersuchungsergebnissen üben die 12 (4*3) Faktorstufenkombinationen überwiegend unterschiedliche Effekte auf die Leistungen der männlichen und weiblichen Testpersonen im EMS-ähnlichen Untertest *Schlauchfiguren* aus. Zudem scheint es,

dass insbesondere die Rate-Bedingung geschlechtsspezifische Unterschiede sichtbar macht.

Die Ergebnisse der multivariaten Kovarianzanalyse verdeutlichen allerdings, dass unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate diese ursprünglich beobachtete signifikante Interaktion höherer Ordnung (Geschlecht x Stereotype Threat x Risikoverhalten) in den Leistungen des EMS-ähnlichen Untertests *Schlauchfiguren* nicht mehr feststellbar ist. Demnach weist das Ergebnis darauf hin, dass durch das Herauspartialisieren des Einflusses der Kovariate (vgl. Bortz & Döring, 2005) Selbstwirksamkeitserwartung, in den Leistungen des EMS-ähnlichen Untertests *Schlauchfiguren*, tatsächlich keine signifikante Interaktion höherer Ordnung angenommen werden kann.

Zudem ließ sich eine zweite Änderung auf die mittels der multivariaten Varianzanalyse gefundenen Ergebnisse feststellen, wenn die Selbstwirksamkeitserwartung als Kovariate berücksichtigt wird. Während sich im Rahmen der multivariaten Varianzanalyse kein signifikanter Unterschied in den Leistungen der EMS-ähnlichen Untertests in Abhängigkeit von den Stereotype Threat-Bedingungen abzeichnete, konnte durch das Herausrechnen des Einflusses der Kovariate (vgl. Bortz & Döring, 2005) Selbstwirksamkeitserwartung, ein signifikanter Unterschied in den Leistungen des EMS-ähnlichen Untertests *Fakten lernen* zwischen den Stereotype Threat-Bedingungen ermittelt werden. Es zeigte sich, dass die Bedingung „Aktivierung“ von Stereotype Threat (vgl. Smith & White, 2002) im Vergleich zu der Bedingung „Neutral“ (es gäbe keine geschlechtsspezifischen Unterschiede in den Leistungen zwischen Männern und Frauen) (vgl. Spencer et al., 1999) einen positiven Effekt auf die Leistungen im EMS-ähnlichen Subtest *Fakten lernen*, der die Merkfähigkeit von bestimmten Fakten erhebt (Hänsgen & Spicher, 2010), auszuüben scheint. Demgemäß konnte erst durch die Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate, verdeutlicht werden, dass tatsächlich ein signifikanter Unterschied in den Leistungen des EMS-ähnlichen Subtests *Fakten lernen* in Abhängigkeit von Stereotype Threat beobachtet werden kann. Die Selbstwirksamkeitserwartung als Kovariate scheint somit bei insgesamt zwei Fragestellungen wirksam zu sein.

Bei Betrachtung der Risikoverhaltens-Tendenzen, die mittels des *Game of Dice Task* (GDT) (Brand et al., 2005) erhoben wurden, konnte erwartungsgemäß festgestellt werden, dass SchülerInnen in der „Risikoverhaltens-Bedingung“ *Raten*, eine höhere Risikoverhaltens-Tendenz im GDT aufweisen im Vergleich zu den Testpersonen in den

beiden anderen Risikoverhaltens-Bedingungen (*sicheres Antworten* und *ohne Instruktion*). An dieser Stelle ist allerdings darauf hinzuweisen, dass sich diese Analyse nicht auf die Gesamtstichprobe (N = 1093), sondern auf eine Teilstichprobe von insgesamt 471 SchülerInnen bezieht, da die Erhebung mittels des GDT nicht bei allen UntersuchungsteilnehmerInnen erfolgen konnte.

In Bezug auf die Geschlechtergruppen zeigen viele Untersuchungen eine höhere Risikoverhaltens-Tendenz zugunsten der Männer auf (z.B.: Charness & Gneezy, 2012). Gemäß den Ergebnissen einer Metaanalyse von Byrnes et al. (1999) konnten geschlechtsspezifische Unterschiede im Risikoverhalten zugunsten der Männer in der Mehrzahl der überprüften Risikoverhaltensbereiche (z.B.: Glücksspiel, Raten) ermittelt werden. Cárdenas et al. (2012) stellten anhand einem Münzwurf, bei dem zwischen einer riskanten oder sicheren Option gewählt werden musste fest, dass männliche Schüler eine höhere Risikoverhaltens-Tendenz aufweisen im Vergleich zu den weiblichen Schülern, wobei dieser Befund kulturübergreifend sowohl in Schweden als auch in Kolumbien beobachtet wurde. Im Gegensatz zur Literatur konnte mit $t(469) = -1.898$, $p = .058$ kein signifikanter Unterschied in der Risikoverhaltens-Tendenz zwischen den männlichen und weiblichen Untersuchungsteilnehmern angenommen werden. Ein etwaiger Grund könnte in dem bereits oben eingegangenen geringeren Stichprobenumfang liegen, der für die Analyse herangezogen werden konnte. Anhand der Berücksichtigung der Mittelwerte, zeichnet sich allerdings, die im Rahmen der Literatur dargelegte höhere Risikoverhaltens-Tendenz zugunsten der Männer (z.B.: Charness & Gneezy, 2012), ab.

Für zukünftige Untersuchungen ist zu empfehlen, die Erfassung der Selbstwirksamkeitserwartung, in Bezug auf die Studie von Marx und Stapel (2006) (im Hinblick auf die Frustration und Ängstlichkeit), zum einen nach den Stereotype Threat Instruktionen gleich im Anschluss an die geschlechtsspezifischen Fähigkeitszuschreibungen und zum anderen nach der Bearbeitung der EMS-ähnlichen Untertests, durchzuführen. Da in der vorliegenden Untersuchung die Selbstwirksamkeitserwartung zum Ende der Testung (erst nach der Theory of Mind und Leistungsmotivation) erhoben wurde, wäre es interessant, ob und inwieweit sich etwaige Unterschiede in den Einschätzungen zur Selbstwirksamkeitserwartung in Abhängigkeit von Geschlecht und den Stereotype Threat Bedingungen im Anschluss an die geschlechtsspezifischen Fähigkeitszuschreibungen und nach der Bearbeitung der EMS-ähnlichen Untertests feststellen lassen. Zudem könnte geklärt werden, ob sich die weiblichen Testpersonen direkt nach den geschlechtsspezifischen

Fähigkeitszuschreibungen selbstwirksamer einschätzen als nach der Bearbeitung der EMS-ähnlichen Untertests.

11. Zusammenfassung

Untersuchungshintergrund stellte der Leistungsunterschied zuungunsten der weiblichen österreichischen Studienbewerber im Vergleich zu den männlichen dar, der sich im Rahmen des EMS (*Eignungstest für das Medizinstudium*) Einsatzes an den Medizinischen Universitäten in Wien und Innsbruck, abzeichnete (Mallinger et al., 2007, 2010). Die vorliegende Arbeit analysierte Faktoren, die auf die Leistungen in EMS-ähnlichen Untertests Einfluss nehmen könnten. In diesem Zusammenhang wurde der Einfluss von Geschlecht, Stereotype Threat und Risikoverhalten auf die Leistungen in EMS-ähnlichen Untertests beleuchtet. Des Weiteren setzte sich die vorliegende Arbeit mit dem Konstrukt der Selbstwirksamkeitserwartung auseinander, hierbei wurden unter anderem geschlechtsspezifische Unterschiede überprüft, ein möglicher Zusammenhang mit den Leistungen in EMS-ähnlichen Untertests analysiert und ein etwaiger Einfluss der Selbstwirksamkeitserwartung als Kovariate untersucht.

Die Stichprobe bestand aus insgesamt 1093 SchülerInnen (62,6 % weiblich und 37,4 % männlich) der 12. Schulstufe von allgemeinbildenden höheren Schulen. Im Rahmen der durchgeführten Testungen wurden sozialstatistische Merkmale, persönliche- und geschlechtsspezifische Fähigkeitszuschreibungen, die Leistungen bei sechs EMS-ähnlichen Untertests, die Theory of Mind, die Leistungsmotivation, die Selbstwirksamkeitserwartung und die Risikoverhaltens-Tendenz, erhoben.

Im Hinblick auf das Konstrukt der Selbstwirksamkeitserwartung zeigten die Ergebnisse einen signifikanten Unterschied in den Einschätzungen zur Selbstwirksamkeitserwartung in Abhängigkeit vom Geschlecht, zugunsten der männlichen Schüler, auf ($H_1(2)$). Die Leistungen in Betracht gezogen, konnte ein schwach positiver, signifikanter Zusammenhang zwischen der Selbstwirksamkeitserwartung und den Leistungen in den EMS-ähnlichen Untertests *Schlauchfiguren* ($r = .100$) und *Diagramme und Tabellen* ($r = .067$), festgestellt werden ($H_1(1)$). Zudem ließ sich ein signifikanter Unterschied in der Selbstwirksamkeitserwartung in Abhängigkeit von der Intention, Medizin zu studieren, erkennen ($H_1(3)$). Hierbei zeichnete sich bei den Testpersonen, die ein Medizinstudium intendieren, eine höhere Selbstwirksamkeitserwartung im Vergleich zu den SchülerInnen die dieses Studium nicht anstreben, ab.

Angesichts der *persönlichen Fähigkeitszuschreibungen* konnten bei den Items zu *Quantitative und formale Probleme*, *Medizinisch-naturwissenschaftlichem Grundverständnis*, *Schlauchfiguren*, *Diagramme und Tabellen* und *Figuren lernen* signifikante Unterschiede in Abhängigkeit vom Geschlecht festgestellt werden: Männliche Schüler wiesen sich bei den fünf angeführten Items höhere Fähigkeiten im Vergleich zu den weiblichen Testpersonen zu.

Die Ergebnisse der multivariaten Varianzanalyse verdeutlichten signifikante Unterschiede in den Leistungen bei fünf EMS-ähnlichen Untertests in Abhängigkeit vom Geschlecht ($H_1(4)$): Männliche Untersuchungsteilnehmer erzielten im Vergleich zu den weiblichen höhere Leistungen bei den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis* sowie *Diagramme und Tabellen*. Im EMS-ähnlichen Untertest *Fakten lernen* konnten Leistungsunterschiede zugunsten der weiblichen Testpersonen beobachtet werden.

Des Weiteren konnten signifikante Wechselwirkungen in den Leistungen bei den EMS-ähnlichen Untertests *Schlauchfiguren* und *Figuren lernen* zwischen Geschlecht und Risikoverhalten (*Raten, ohne Instruktion* und *sicheres Antworten*) ermittelt werden ($H_1(8)$). Die Rate-Bedingung scheint die Leistungen der männlichen Schüler im Vergleich zu den weiblichen Testpersonen zu begünstigen.

Im Hinblick auf den Faktor *Stereotype Threat* konnten die Ergebnisse der *geschlechtsspezifischen Fähigkeitszuschreibungen* die Wirksamkeit der *Stereotype Threat*-Bedingungen (*Aktivierung*, *neutrale Bedingung*, *Aufklärung* und *Negativ-Aktivierung*) untermauern. Allerdings ließ sich keine signifikante Wechselwirkung in den Leistungen bei den EMS-ähnlichen Untertests zwischen Geschlecht und *Stereotype Threat* feststellen ($H_1(7)$).

In Bezug auf die Risikoverhaltens-Tendenz zeigte sich, dass Untersuchungsteilnehmer in der *Rate*-Bedingung eine höhere Risikoverhaltens-Tendenz im *Game of Dice Task* (Brand et al., 2005) aufwiesen als SchülerInnen in den anderen Bedingungen (*sicheres Antworten* und *ohne Instruktion*) ($H_1(18)$). Während sich anhand der Mittelwerte eine höhere Risikoverhaltens-Tendenz zugunsten der männlichen Schüler abzeichnete, konnte kein signifikanter Unterschied in der Risikoverhaltens-Tendenz zwischen den männlichen und weiblichen Testpersonen verdeutlicht werden ($H_1(19)$).

Die Prüfung einer signifikanten Interaktion höherer Ordnung aus Geschlecht x *Stereotype Threat* x Risikoverhalten in den Leistungen der EMS-ähnlichen Untertests,

bestätigte sich im Subtest *Schlauchfiguren* ($H_1(10)$). Jedoch wurde dieser Befund unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate widerlegt ($H_1(17)$). Daneben ergab sich noch eine weitere Änderung der ursprünglich beobachteten Resultate der multivariaten Varianzanalyse. Unter Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate konnte nun ein signifikanter Unterschied in den Leistungen des EMS-ähnlichen Untertests *Fakten lernen* zwischen den Stereotype Threat-Bedingungen aufgezeigt werden ($H_1(12)$). Die Ergebnisse deuten darauf hin, dass die Stereotype Threat-Bedingung „Aktiverung“ im Vergleich zur Bedingung „Neutral“ die Leistungen in diesem Untertest zu begünstigen scheint.

12. Abstract

Untersucht wurden der Einfluss von Geschlecht, Stereotype Threat-Bedingungen und Risikoverhalten-Bedingungen auf die Leistungen bei insgesamt sechs EMS-ähnlichen Untertests. Hintergrund stellte der im EMS (*Eignungstest für das Medizinstudium*) aufgezeigte Leistungsunterschied zuungunsten der weiblichen Studienbewerber aus Österreich, im Vergleich mit den männlichen, dar (Mallinger et al., 2007, 2010). Ein weiterer Fokus lag auf dem Effekt der Selbstwirksamkeitserwartung; neben der Analyse als potentielle Kovariate wurden mögliche geschlechtsspezifische Unterschiede sowie ein Zusammenhang mit den Leistungen beleuchtet.

Eine Stichprobe von insgesamt 1093 SchülerInnen (62,6 % weiblich und 37,4 % männlich) mit einem Durchschnittsalter von 17.57 Jahren (± 0.76) der 12. Schulstufe allgemeinbildender höherer Schulen lag den Untersuchungen zugrunde. Im Rahmen der Erfassung der Risikoverhaltens-Tendenz, mittels *Game of Dice Task* (Brand et al., 2005), konnte eine Teilstichprobe von 471 SchülerInnen gewonnen werden.

Geschlechtsspezifische Leistungsunterschiede ließen sich bei den EMS-ähnlichen Untertests *Quantitative und formale Probleme*, *Schlauchfiguren*, *Medizinisch-naturwissenschaftliches Grundverständnis* und *Diagramme und Tabellen* zugunsten der männlichen und bei *Fakten lernen* zugunsten der weiblichen Schüler erkennen. Männliche Testpersonen schrieben sich eine höhere Selbstwirksamkeitserwartung zu als weibliche. Es konnten schwach positive, signifikante Zusammenhänge zwischen der Selbstwirksamkeitserwartung und Leistungen in den EMS-ähnlichen Untertests *Schlauchfiguren* sowie *Diagramme und Tabellen* beobachtet werden.

Im Hinblick auf den Faktor Risikoverhalten konnte die Instruktion zum Raten (vgl. Ben-Shakhar & Sinai, 1991) bei zwei EMS-ähnlichen Untertests die Leistungen der männlichen Schüler im Vergleich zu den weiblichen positiv beeinflussen. Während die Wirksamkeit der Stereotype Threat-Bedingungen belegt werden konnte, ließ sich jedoch keine signifikante Wechselwirkung in den Leistungen der EMS-ähnlichen Untertests zwischen Geschlecht und Stereotype Threat beobachten. Ein signifikanter Unterschied in der Risikoverhaltens-Tendenz zwischen den männlichen und weiblichen Testpersonen konnte nicht beobachtet werden. Die Berücksichtigung der Selbstwirksamkeitserwartung als Kovariate zeigte zwei Änderungen der ursprünglich gefundenen Ergebnisse der multivariaten Varianzanalyse auf: Es konnte nun ein

signifikanter Unterschied in den Leistungen des EMS-ähnlichen Untertests *Fakten lernen* zwischen den Stereotype Threat-Bedingungen verdeutlicht werden sowie zeichnete sich keine signifikante Interaktion höherer Ordnung aus Geschlecht x Stereotype Threat x Risikoverhalten in den Leistungen des EMS-ähnlichen Untertests *Schlauchfiguren* ab.

13. Literaturverzeichnis

- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M. & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29-46.
- Arthur, N. & Everaert, P. (2012). Gender and performance in accounting examinations: Exploring the impact of examination format. *Accounting Education: an international journal*, 21 (5), 471-487.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2011). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (13. Aufl.). Berlin: Springer.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. New Jersey: Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Ben-Shakhar, G. & Sinai, Y. (1991). Gender differences in multiple-choice tests: The role of differential guessing tendencies. *Journal of Educational Measurement*, 28 (1), 23-35.
- Bolger, N. & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*, 27 (2), 165-174.
- Bortz, J. & Döring, N. (2005). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. 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 (3), 267-277.
- Bühl, A. (2006). *SPSS Version 14. Einführung in die moderne Datenanalyse* (10., überarbeitete Aufl.). München: Pearson Studium.

- Bundesministerium für Wissenschaft und Forschung (BMWF). (2002). Universitätsgesetz 2002 - Österreichisches Hochschulrecht (Stand: BGBl. I Nr. 13/2011) Bundesgesetz über die Organisation der Universitäten und ihre Studien (Universitätsgesetz 2002 - UG) [online]. URL: http://www.bmwf.gv.at/uploads/tx_contentbox/UG-2002_BGBl.I_Nr._13-2011.pdf [08.09.11].
- Byrnes, J. P., Miller, D. C. & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125 (3), 367-383.
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L. & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, 33, 267-285.
- Cadinu, M., Maass, A., Rosabianca, A. & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, 16, 572-578.
- Campbell, S. M. & Collaer, M. L. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychology of Women Quarterly*, 33, 437-444.
- Cárdenas, J.-C., Dreber, A., von Essen, E. & Ranehill, E. (2012). Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior & Organization*, 83, 11-23.
- Charness, G. & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83, 50-58.
- Chung, B. G., Ehrhart, M. G., Holcombe Ehrhart, K., Hattrup, K. & Solamon, J. (2010). Stereotype threat, state anxiety, and specific self-efficacy as predictors of promotion exam performance. *Group & Organization Management*, 35 (1), 77-107.
- Edelstein, W., Hopf, D., Jerusalem, M. & Schwarzer, R. (1997). Wissenschaftliche Begleitung: Aufbau, Arbeitsansätze, erste Ergebnisse. In R. Brockmeyer & W. Edelstein (Hrsg.), *Selbstwirksame Schulen: Wege pädagogischer Innovation* (S. 135-152). Oberhausen: Karl Maria Laufen.
- Fiedler, K. & Bless, H. (2002). Soziale Kognition. In W. Stroebe, K. Jonas & M. Hewstone (Hrsg.), *Sozialpsychologie: Eine Einführung* (4. Aufl.) (S. 125-163). Berlin: Springer.

- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.
- Good, C., Aronson, J. & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*, 645-662.
- Halpern, D. F. & Tan, U. (2001). Stereotypes and steroids: Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cognition, 45*, 392-414.
- Hanna, G. (1986). Sex differences in the mathematics achievement of eight graders in Ontario. *Journal for Research in Mathematics Education, 17* (3), 231-237.
- Hänsgen, K.-D. (2007). Numerus clausus in der Medizin – werden die Richtigen ausgewählt für Studium und Beruf? *Schweizerische Ärztezeitung, 88* (46), 1953-1957.
- Hänsgen, K.-D. & Spicher, B. (2000). Erfahrungen mit dem EMS als Zulassungskriterium: Zwei Jahre Numerus clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 1. *Schweizerische Ärztezeitung, 81* (13), 666-672.
- Hänsgen, K.-D. & Spicher, B. (2001). *EMS - Eignungstest für das Medizinstudium in der Schweiz 2001: Bericht über die Durchführung und Ergebnisse 2001. Vorhersage des Prüfungserfolges - Trends und Schlussfolgerungen aus vier Jahren Testanwendung*. Freiburg: Zentrum für Testentwicklung und Diagnostik.
- Hänsgen, K.-D. & Spicher, B. (2010). *EMS - Eignungstest für das Medizinstudium 2010: Bericht 17 über die Durchführung und Ergebnisse 2010*. Freiburg: Zentrum für Testentwicklung und Diagnostik.
- Hinz, A., Schumacher, J., Albani, C., Schmid, G. & Brähler, E. (2006). Bevölkerungsrepräsentative Normierung der Skala zur Allgemeinen Selbstwirksamkeitserwartung. *Diagnostica, 52* (1), 26-32.
- Hofmann, K. T. (1997). Aufbau, Organisation und Steuerung. In R. Brockmeyer & W. Edelstein (Hrsg.), *Selbstwirksame Schulen: Wege pädagogischer Innovation* (S. 33-38). Oberhausen: Karl Maria Laufen.
- ITB Consulting (Hrsg.). (2008). *Test für medizinische Studiengänge II: Originalversion II des TMS* (5. Aufl.). Göttingen: Hogrefe.

- Jerusalem, M. (2002). Einleitung. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, Zeitschrift für Pädagogik, 44. Beiheft* (S. 8-12). Weinheim: Beltz.
- Jerusalem, M. & Mittag, W. (1999). Selbstwirksamkeit, Bezugsnormen, Leistung und Wohlbefinden in der Schule. In M. Jerusalem & R. Pekrun (Hrsg.), *Emotion, Motivation und Leistung* (S. 223-245). Göttingen: Hogrefe.
- Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung (WIRKSCHUL). In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (S. 15-16). Berlin: Freie Universität Berlin.
- Jerusalem, M. & Schwarzer, R. (1999). Allgemeine Selbstwirksamkeitserwartung (WIRKALL_r). In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (S. 13-14). Berlin: Freie Universität Berlin.
- 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.
- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female student's maths performance. *British Journal of Educational Psychology, 77*, 323-338.
- Keller, J. & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin, 29*, 371-381.
- Lent, R. W., Brown, S. D. & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology, 31* (3), 356-362.
- Lent, R. W., Brown, S. D. & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology, 33* (3), 265-269.

- Luszczynska, A., Gutiérrez-Doña, B. & Schwarzer, R. (2005). General self-efficacy in various domains of human functioning: Evidence from five countries. *International Journal of Psychology*, 40 (2), 80-89.
- Mallinger, R., Holzbaur, C., Dierich, M., Heidegger, M., Hänsgen, K.-D. & Spicher, B. (2007). *EMS - Eignungstest für das Medizinstudium in Österreich - Medizinische Universität Innsbruck, Medizinische Universität Wien 2007*. Freiburg: Zentrum für Testentwicklung und Diagnostik.
- Mallinger, R., Holzbaur, C., Dierich, M., Heidegger, M., Hänsgen, K.-D. & Spicher, B. (2008). *EMS - Eignungstest für das Medizinstudium in Österreich - Medizinische Universität Innsbruck, Medizinische Universität Wien 2008*. Freiburg: Zentrum für Testentwicklung und Diagnostik.
- Mallinger, R., Holzbaur, C., Mutz, N., Prodinger, W. M., Heidegger, M., Hänsgen, K.-D., Spicher, B., Guntern, S. & Meissner, E. (2010). *EMS - Eignungstest für das Medizinstudium in Österreich mit einem Bericht über die Pilotstudie zum Self-Assessment „Studienrelevante Persönlichkeitsmerkmale“ - Medizinische Universität Innsbruck, Medizinische Universität Wien 2010*. Freiburg: Zentrum für Testentwicklung und Diagnostik.
- Marx, D. M. & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, 28, 1183-1193.
- 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.
- McIntyre, R. B., Paulson, R. M. & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, 39, 83-90.
- Medizinische Universität Wien (MedUni Wien). (2012a). 4.352 TeilnehmerInnen beim Eignungstest für das Medizinstudium an der MedUni Wien [online]. URL: [http://www.meduniwien.ac.at/homepage/news-und-topstories/?tx_ttnews\[tt_news\]=2799&cHash=9b3e3671d1](http://www.meduniwien.ac.at/homepage/news-und-topstories/?tx_ttnews[tt_news]=2799&cHash=9b3e3671d1) [18.12.12].
- Medizinische Universität Wien (MedUni Wien). (2012b). Neue Aufnahmetests für Medizinstudium ab 2013 [online]. URL: [http://www.meduniwien.ac.at/homepage/news-und-topstories/?tx_ttnews\[tt_news\]=3142&cHash=9e170ba8b1](http://www.meduniwien.ac.at/homepage/news-und-topstories/?tx_ttnews[tt_news]=3142&cHash=9e170ba8b1) [18.12.12].

- Mittag, W., Kleine, D. & Jerusalem, M. (2002). Evaluation der schulbezogenen Selbstwirksamkeit von Sekundarschülern. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, Zeitschrift für Pädagogik, 44. Beiheft* (S. 145-173). Weinheim: Beltz.
- Moritz, S., Veckenstedt, R., Randjbar, S. & Vitzthum, F. (2011). *MKT+. Individualisiertes Metakognitives Therapieprogramm für Menschen mit Psychose*. Berlin: Springer.
- Multon, K. D., Brown, S. D. & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38* (1), 30-38.
- Oswald, D. L. & Harvey, R. D. (2000). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social, 19* (4), 338-356.
- Pajares, F. & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426-443.
- Pajares, F. & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86* (2), 193-203.
- Satow, L. & Mittag, W. (1999). Selbstwirksamkeitserwartung im Umgang mit sozialen Anforderungen (WIRKSOZ). In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (S. 17-18). Berlin: Freie Universität Berlin.
- Schuler, H. & Prochaska, M. (2001). *Leistungsmotivationsinventar (LMI)*. Göttingen: Hogrefe.
- Schwarzer, R. (2002). SWE. Skala zur Allgemeinen Selbstwirksamkeitserwartung. In E. Brähler, J. Schumacher & B. Strauß (Hrsg.), *Diagnostische Verfahren in der Psychotherapie (Diagnostik für Klinik und Praxis, Bd. 1)* (S. 362-365). Göttingen: Hogrefe.
- Schwarzer, R. (2004). *Psychologie des Gesundheitsverhaltens. Einführung in die Gesundheitspsychologie* (3., überarbeitete Aufl.). Göttingen: Hogrefe.

- Schwarzer, R. (2009). Everything you wanted to know about the general self-efficacy scale but were afraid to ask [online]. URL: http://userpage.fu-berlin.de/~health/faq_gse.pdf [24.03.11].
- Schwarzer, R., Bäßler, J., Kwiatek, P., Schröder, K. & Zhang, J. X. (1997). The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the General Self-efficacy Scale. *Applied Psychology: An International Review*, 46 (1), 69-88.
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, Zeitschrift für Pädagogik*, 44. Beiheft (S. 28-53). Weinheim: Beltz.
- Schwarzer, R., Mueller, J. & Greenglass, E. (1999). Assessment of perceived general self-efficacy on the internet: Data collection in cyberspace. *Anxiety, Stress, and Coping*, 12, 145-161.
- Sekaquaptewa, D. & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effect on women's performance. *Journal of Experimental Social Psychology*, 39, 68-74.
- Smith, J. L. & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles*, 47, 179-191.
- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Spiel, C., Schober, B. & Litzenberger, M. (2008). *Evaluation der Eignungstests für das Medizinstudium in Österreich* (Projektbericht). Wien: Universität, Fakultät für Psychologie.
- Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *American Psychologist*, 52 (6), 613-629.
- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69 (5), 797-811.
- Stumpf, H. & Fay, E. (1983). *Schlauchfiguren - Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens*. Göttingen: Hogrefe.

- Trimpop, R. M. (1994). *The psychology of risk-taking behavior*. Amsterdam: North Holland, Elsevier Science.
- Von Schrader, S. & Ansley, T. (2006). Sex differences in the tendency to omit items on multiple-choice tests: 1980-2000. *Applied Measurement in Education*, 19 (1), 41-65.
- Willinger, U., Schmöger, M., Müller, C. & Auff, E. (in Vorbereitung). *Theory of mind stories*.
- Zentrum für Testentwicklung und Diagnostik (ZTD). (2009). *TestInfo '09. Eignungstest für das Medizinstudium (EMS). Version A*. [Broschüre]. Universität Freiburg und die Medizinischen Universitäten Innsbruck und Wien.

14. Abbildungsverzeichnis

Abbildung 1: Abgrenzung zwischen Selbstwirksamkeitserwartungen und Konsequenzerwartungen (Bandura, 1977, S. 193)	19
Abbildung 2: Itembeispiel geschlechtsspezifische Fähigkeitszuschreibung; Untertest Medizinisch-naturwissenschaftliches Grundverständnis (Hänsgen & Spicher, 2010)	41
Abbildung 3: Itembeispiel persönliche Fähigkeitszuschreibung; Untertest Figuren lernen (Hänsgen & Spicher, 2010).....	42
Abbildung 4: Itembeispiel Quantitative und formale Probleme (Hänsgen & Spicher, 2010, S. 93).....	43
Abbildung 5: Itembeispiel Schlauchfiguren (Hänsgen & Spicher, 2010, S. 94)	44
Abbildung 6: Itembeispiel Medizinisch-naturwissenschaftliches Grundverständnis (Hänsgen & Spicher, 2010, S. 97f.)	45
Abbildung 7: Itembeispiel Diagramme und Tabellen (Hänsgen & Spicher, 2010, S. 100)	46
Abbildung 8: Itembeispiel Figuren lernen (Hänsgen & Spicher, 2010, S. 99)	47
Abbildung 9: Itembeispiel Fakten lernen (Hänsgen & Spicher, 2010, S. 99)	47
Abbildung 10: Itembeispiel Leistungsmotivationsinventar (Schuler & Prochaska, 2001)	49
Abbildung 11: Itembeispiel Skala zur Allgemeinen Selbstwirksamkeitserwartung (Jerusalem & Schwarzer, 1999).....	50
Abbildung 12: Game of Dice Task (Brand et al., 2005, S. 270).....	51
Abbildung 13: Geschlecht (N = 1093 gültige Fälle).....	56
Abbildung 14: Staatsbürgerschaft (N = 1050 gültige Fälle)	58
Abbildung 15: Ausbildungszweig (N = 1080 gültige Fälle)	60
Abbildung 16: Höchste abgeschlossene Ausbildung der Mütter (N = 1063 gültige Fälle)	63
Abbildung 17: Höchste abgeschlossene Ausbildung der Väter (N = 1048 gültige Fälle)	65

Abbildung 18: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Quantitative und formale Probleme	79
Abbildung 19: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Schlauchfiguren	81
Abbildung 20: Relative Lösungshäufigkeit im EMS-ähnlichen Untertest Medizinisch-naturwissenschaftliches Grundverständnis.....	82
Abbildung 21: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Diagramme und Tabellen	83
Abbildung 22: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Figuren lernen	84
Abbildung 23: relative Lösungshäufigkeit im EMS-ähnlichen Untertest Fakten lernen .	85
Abbildung 24: Histogramm zum Meanscore SWE	87
Abbildung 25: Histogramm zum Nettoscore-GDT	88
Abbildung 26: Wechselwirkung Geschlecht * Stereotype Threat bei persönlicher Fähigkeitszuschreibung Diagramme und Tabellen	100
Abbildung 27: Wechselwirkung Geschlecht * Risikoverhalten im UT Schlauchfiguren	110
Abbildung 28: Wechselwirkung Geschlecht * Risikoverhalten im UT Figuren lernen .	111
Abbildung 29: Wechselwirkung Stereotype Threat * Risikoverhalten im UT Diagramme und Tabellen	112
Abbildung 30: Interaktion höherer Ordnung im UT Schlauchfiguren, männliche Testpersonen	113
Abbildung 31: Interaktion höherer Ordnung im UT Schlauchfiguren, weibliche Testpersonen	113
Abbildung 32: Boxplot Diagramme GDT	121

15. Tabellenverzeichnis

Tabelle 1: EMS-Teststruktur (nach Hänsgen & Spicher, 2010, S. 41).....	4
Tabelle 2: geplantes Studiendesign.....	38
Tabelle 3: Erhebungsinstrumente.....	40
Tabelle 4: EMS-ähnliche Untertests in der vorliegenden Studie, Anzahl der Items und Bearbeitungszeit (nach Hänsgen & Spicher, 2010, S. 41).....	42
Tabelle 5: interne Konsistenz (Cronbach´s Alpha) der Untertests.....	48
Tabelle 6: Chi-Quadrat-Test Geschlecht	56
Tabelle 7: Häufigkeitsverteilung Alter (N = 1086 gültige Fälle).....	57
Tabelle 8: Chi-Quadrat-Test Alter	57
Tabelle 9: Chi-Quadrat-Test Staatsbürgerschaft.....	58
Tabelle 10: Häufigkeitsverteilung Ausbildungszweig (N = 1080 gültige Fälle).....	59
Tabelle 11: Chi-Quadrat-Test Ausbildungszweig	60
Tabelle 12: Häufigkeitsverteilung Wiederholung einer Schulstufe (N = 1087 gültige Fälle).....	61
Tabelle 13: Chi-Quadrat-Test Wiederholung einer Schulstufe	61
Tabelle 14: Kontingenztafel Geschlecht und Wiederholung zumindest einer Schulstufe	62
Tabelle 15: Häufigkeitsverteilung höchste abgeschlossenen Ausbildung der Mütter (N = 1063 gültige Fälle)	63
Tabelle 16: Chi-Quadrat-Test höchste abgeschlossene Ausbildung der Mütter	64
Tabelle 17: Häufigkeitsverteilung höchste abgeschlossene Ausbildung der Väter (N = 1048 gültige Fälle)	64
Tabelle 18: Chi-Quadrat-Test höchste abgeschlossene Ausbildung der Väter	65
Tabelle 19: Häufigkeitsverteilung Intention, Medizin zu studieren.....	66
Tabelle 20: Chi-Quadrat-Test Intention, Medizin zu studieren	66
Tabelle 21: Kontingenztafel Geschlecht und die Intention, Medizin zu studieren	67

Tabelle 22: Kontingenztabelle Geschlecht und die Befassung mit dem EMS	68
Tabelle 23: Kontingenztabelle die Intention, Medizin zu studieren und die Befassung mit dem EMS	69
Tabelle 24: Zuweisung der UntersuchungsteilnehmerInnen zu den Stereotype Threat - Bedingungen.....	70
Tabelle 25: Chi-Quadrat-Test Stereotype Threat-Bedingungen.....	70
Tabelle 26: Zuweisung der UntersuchungsteilnehmerInnen zu den Risikoverhalten - Bedingungen.....	70
Tabelle 27: Chi-Quadrat-Test Risikoverhalten-Bedingungen	71
Tabelle 28: Kontingenztabelle Geschlecht und Stereotype Threat - Bedingungen	71
Tabelle 29: Kontingenztabelle Geschlecht und Risikoverhalten - Bedingungen.....	72
Tabelle 30: Kontingenztabelle Stereotype Threat - Bedingungen und Risikoverhalten - Bedingungen.....	73
Tabelle 31: Deskriptivstatistik EMS-ähnliche Subtests (N = 1093).....	74
Tabelle 32: Reliabilitätskoeffizienten (Cronbach´s Alpha) der EMS-ähnlichen Untertests	75
Tabelle 33: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang, Itemanzahl) EMS-ähnlich. Subtest Quantitative und formale Probleme	75
Tabelle 34: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnlich. Subtest Schlauchfiguren	76
Tabelle 35: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnlich. Subtest Medizinisch-naturwissenschaftliches Grundverständnis	76
Tabelle 36: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnlich. Subtest Diagramme und Tabellen.....	77
Tabelle 37: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnlich. Subtest Figuren lernen.....	77
Tabelle 38: Reliabilitätskoeffizienten (Cronbach´s Alpha) der Faktorstufenkombinationen (Stichprobenumfang; Itemanzahl) EMS-ähnlich. Subtest Fakten lernen.....	77

Tabelle 39: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Quantitative und formale Probleme	78
Tabelle 40: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Schlauchfiguren	80
Tabelle 41: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Medizinisch-naturwissenschaftliches Grundverständnis	81
Tabelle 42: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Diagramme und Tabellen	82
Tabelle 43: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Figuren lernen	84
Tabelle 44: Lösungswahrscheinlichkeiten und Trennschärfen im EMS-ähnlichen Subtest Fakten lernen	85
Tabelle 45: Itemstatistik Selbstwirksamkeitserwartung (N = 1034 gültige Fälle)	86
Tabelle 46: Deskriptivstatistik Selbstwirksamkeitserwartung	87
Tabelle 47: Deskriptivstatistik Game of Dice Task	88
Tabelle 48: Häufigkeitsverteilung der Items zur geschlechtsspezifischen Fähigkeitszuschreibung in Abhängigkeit von der Wirksamkeit der Instruktionen	90
Tabelle 49: Chi-Quadrat-Tests Items zur geschlechtsspezifischen Fähigkeitszuschreibung	90
Tabelle 50: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Quantitative und formale Probleme	91
Tabelle 51: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Schlauchfiguren	92
Tabelle 52: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Medizinisch-naturwissenschaftliches Grundverständnis .	93
Tabelle 53: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Diagramme und Tabellen	94
Tabelle 54: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Figuren lernen	95
Tabelle 55: Kontingenztafel Stereotype Threat-Bedingungen und Wirksamkeit der Instruktionen bei Fakten lernen	96
Tabelle 56: Levene´s Test, persönliche Fähigkeitszuschreibungen	97
Tabelle 57: Tests der Zwischensubjekteffekte, Quelle: Geschlecht, AV: persönl. Fähigkeitszuschreibung	98

Tabelle 58: Deskriptivstatistik persönliche Fähigkeitszuschreibungen in Abhängigkeit vom Geschlecht (Frauen: N = 670, Männer: N = 395).....	98
Tabelle 59: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat, AV: persönl. Fähigkeitszuschreibung	99
Tabelle 60: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat, AV: persönl. Fähigkeitszuschreibung	99
Tabelle 61: Produkt-Moment-Korrelation, SWE mit den EMS-ähnlichen Untertests (N = 1071)	101
Tabelle 62: Produkt-Moment-Korrelation, SWE mit den EMS-ähnlichen Untertests in Abhängigkeit vom Geschlecht (Frauen: N=673, Männer: N = 398)	102
Tabelle 63: Tests der Zwischensubjekteffekte, AV: Selbstwirksamkeitserwartung.....	103
Tabelle 64: Deskriptivstatistik SWE-Meanscore in Abhängigkeit vom Geschlecht	103
Tabelle 65: t-Test bei unabhängigen Stichproben, SWE und Intention, Medizin zu studieren.....	104
Tabelle 66: Gruppenstatistik SWE in Abhängigkeit von der Intention, Medizin zu studieren.....	105
Tabelle 67: Levene´s Test, EMS-ähnliche Untertests	106
Tabelle 68: Tests der Zwischensubjekteffekte, Quelle: Geschlecht	106
Tabelle 69: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit vom Geschlecht (Frauen: N = 684, Männer: N = 409).....	107
Tabelle 70: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat.....	107
Tabelle 71: Tests der Zwischensubjekteffekte, Quelle: Risikoverhalten.....	108
Tabelle 72: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit von den Risikoverhalten-Bedingungen	109
Tabelle 73: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat	109
Tabelle 74: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Risikoverhalten	110
Tabelle 75: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat * Risikoverhalten	111
Tabelle 76: Tests der Zwischensubjekteffekte, Quelle: Geschlecht* Stereotype Threat * Risikoverhalten	112
Tabelle 77: Levene´s Test, EMS-ähnliche Untertests (unter Berücksichtigung SWE)	114

Tabelle 78: Tests der Zwischensubjekteffekte, Quelle: Selbstwirksamkeitserwartung	115
Tabelle 79: Tests der Zwischensubjekteffekte, Quelle: Geschlecht, Kovariate SWE	115
Tabelle 80: Deskriptivstatistik EMS-ähnliche Untertests in Abhängigkeit vom Geschlecht unter Berücksichtigung SWE (Frauen: N = 673, Männer: N = 398)	116
Tabelle 81: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat, Kovariate SWE	117
Tabelle 82: Tests der Zwischensubjekteffekte, Quelle: Risikoverhalten, Kovariate SWE	117
Tabelle 83: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Stereotype Threat, Kovariate SWE.....	118
Tabelle 84: Tests der Zwischensubjekteffekte, Quelle: Geschlecht * Risikoverhalten, Kovariate SWE.....	118
Tabelle 85: Tests der Zwischensubjekteffekte, Quelle: Stereotype Threat * Risikoverhalten, Kovariate SWE.....	119
Tabelle 86: Tests der Zwischensubjekteffekte, Quelle: Geschlecht *Stereotype Threat * Risikoverhalten, Kovariate SWE.....	119
Tabelle 87: Deskriptivstatistik GDT (Gesamt und Risikoverhalten-Bedingungen)	120
Tabelle 88: Prüfgrößen und Parameter zum Gruppenvergleich mittels a-priori Kontrasten, GDT	121
Tabelle 89: Gruppenstatistik NS-GDT	122
Tabelle 90: t-Test bei unabhängigen Stichproben, Risikoverhaltens-Tendenz und Geschlecht.....	122

16. Anhang

Deskriptivstatistik

A) geschlechtsspezifische Fähigkeitszuschreibungen

(Mittelwerte der Prozentangaben zu den sechs Items)

Stereotype Threat-Bedingungen	Geschlechtsspezifische Fähigkeitszuschreibung	N	M	SD
Aktivierung	Figuren lernen - Männlich	289	42,78	24,143
	Figuren lernen - Weiblich	289	57,25	24,122
	Fakten lernen - Männlich	290	33,69	23,036
	Fakten lernen - Weiblich	290	66,22	23,196
	Med.nat. Grundv. - Männlich	290	59,21	22,244
	Med.nat. Grundv. - Weiblich	290	40,76	22,257
	Diagramme u. Tab.- Männlich	289	69,71	19,777
	Diagramme u. Tab.- Weiblich	289	30,35	19,829
	Quant. u. formale Pr.- Männlich	289	68,56	21,207
	Quant. u. formale Pr. - Weiblich	289	31,54	21,322
	Schlauchfiguren- Männlich	289	71,65	20,886
	Schlauchfiguren - Weiblich	289	28,45	21,023
	Gültige Werte (Listenweise)	286		
Neutral	Figuren lernen - Männlich	255	49,47	8,000
	Figuren lernen - Weiblich	255	50,76	7,981
	Fakten lernen - Männlich	254	48,13	13,263
	Fakten lernen - Weiblich	254	51,87	13,263
	Med.nat. Grundv. - Männlich	254	52,78	9,611
	Med.nat. Grundv. - Weiblich	254	47,26	9,585
	Diagramme u. Tab.- Männlich	254	54,57	9,236
	Diagramme u. Tab.- Weiblich	254	45,44	9,238
	Quant. u. formale Pr.- Männlich	254	53,70	12,333
	Quant. u. formale Pr. - Weiblich	254	46,10	11,863
	Schlauchfiguren- Männlich	254	53,08	12,247
	Schlauchfiguren - Weiblich	254	47,00	12,137
	Gültige Werte (Listenweise)	254		

Aufklärung	Figuren lernen - Männlich	279	50,06	8,606
	Figuren lernen - Weiblich	279	49,90	8,730
	Fakten lernen - Männlich	278	46,98	14,125
	Fakten lernen - Weiblich	278	53,00	14,221
	Med.nat. Grundv. - Männlich	277	52,12	9,706
	Med.nat. Grundv. - Weiblich	277	47,88	9,668
	Diagramme u. Tab.- Männlich	277	55,09	11,511
	Diagramme u. Tab.- Weiblich	277	44,84	11,142
	Quant. u. formale Pr.- Männlich	277	56,45	11,803
	Quant. u. formale Pr. - Weiblich	277	43,58	11,810
	Schlauchfiguren- Männlich	277	54,62	12,270
	Schlauchfiguren - Weiblich	277	45,36	12,274
	Gültige Werte (Listenweise)	274		
Negativ-Aktivierung	Figuren lernen - Männlich	261	55,95	25,291
	Figuren lernen - Weiblich	261	44,09	25,308
	Fakten lernen - Männlich	260	61,13	26,790
	Fakten lernen - Weiblich	260	38,60	26,827
	Med.nat. Grundv. - Männlich	259	41,60	24,953
	Med.nat. Grundv. - Weiblich	260	58,22	25,085
	Diagramme u. Tab.- Männlich	260	36,09	25,020
	Diagramme u. Tab.- Weiblich	260	64,12	24,786
	Quant. u. formale Pr.- Männlich	260	42,84	26,155
	Quant. u. formale Pr. - Weiblich	260	57,06	26,378
	Schlauchfiguren- Männlich	260	36,35	26,767
	Schlauchfiguren - Weiblich	260	63,89	26,499
	Gültige Werte (Listenweise)	259		

B) persönliche Fähigkeitszuschreibungen

Persönl. Fähigkeitszuschr.	Geschlecht	Stereotype Threat-Bedingungen	<i>M</i>	<i>SD</i>	<i>N</i>
Figuren lernen	männlich	Aktivierung	6,59	,775	119
		Neutral	6,73	,664	83
		Aufklärung	6,55	1,106	108
		Negativ-Aktivierung	6,67	,836	85
		Gesamt	6,63	,870	395
	weiblich	Aktivierung	6,55	1,049	168
		Neutral	6,51	1,000	164
		Aufklärung	6,53	1,033	165
Negativ-Aktivierung		6,41	1,083	173	

	Gesamt	Gesamt	6,50	1,041	670
		Aktivierung	6,56	,943	287
		Neutral	6,58	,906	247
		Aufklärung	6,54	1,060	273
		Negativ-Aktivierung	6,50	1,014	258
		Gesamt	6,55	,983	1065
Fakten lernen	männlich	Aktivierung	4,84	1,275	119
		Neutral	5,16	1,418	83
		Aufklärung	4,81	1,284	108
		Negativ-Aktivierung	4,98	1,215	85
		Gesamt	4,93	1,298	395
	weiblich	Aktivierung	5,05	1,210	168
		Neutral	5,17	1,170	164
		Aufklärung	5,10	1,213	165
		Negativ-Aktivierung	4,88	1,293	173
		Gesamt	5,05	1,225	670
	Gesamt	Aktivierung	4,97	1,240	287
		Neutral	5,17	1,256	247
		Aufklärung	4,99	1,247	273
		Negativ-Aktivierung	4,91	1,266	258
		Gesamt	5,01	1,254	1065
Med. nat. Grundv.	männlich	Aktivierung	5,30	1,344	119
		Neutral	5,10	1,385	83
		Aufklärung	5,05	1,469	108
		Negativ-Aktivierung	4,84	1,511	85
		Gesamt	5,09	1,428	395
	weiblich	Aktivierung	4,82	1,445	168
		Neutral	4,93	1,451	164
		Aufklärung	4,82	1,402	165
		Negativ-Aktivierung	4,89	1,260	173
		Gesamt	4,86	1,388	670
	Gesamt	Aktivierung	5,02	1,421	287
		Neutral	4,98	1,428	247
		Aufklärung	4,91	1,431	273
		Negativ-Aktivierung	4,87	1,345	258
		Gesamt	4,95	1,407	1065
Diagram me u. Tabellen	männlich	Aktivierung	5,26	1,498	119
		Neutral	5,42	1,458	83
		Aufklärung	5,07	1,569	108

	weiblich	Negativ-Aktivierung	4,56	1,644	85	
		Gesamt	5,09	1,565	395	
		Aktivierung	3,96	1,615	168	
		Neutral	4,23	1,706	164	
		Aufklärung	4,24	1,418	165	
		Negativ-Aktivierung	4,45	1,612	173	
	Gesamt	Gesamt	4,22	1,597	670	
		Aktivierung	4,50	1,691	287	
		Neutral	4,63	1,720	247	
		Aufklärung	4,57	1,533	273	
		Negativ-Aktivierung	4,48	1,620	258	
	Quantitative u. formale Probleme	männlich	Gesamt	4,54	1,640	1065
			Aktivierung	5,43	1,499	119
			Neutral	5,00	1,753	83
			Aufklärung	4,93	1,700	108
Negativ-Aktivierung			4,74	1,767	85	
weiblich		Gesamt	5,05	1,683	395	
		Aktivierung	3,99	1,771	168	
		Neutral	4,06	1,846	164	
		Aufklärung	3,98	1,737	165	
		Negativ-Aktivierung	4,12	1,801	173	
Gesamt		Gesamt	4,04	1,786	670	
		Aktivierung	4,59	1,807	287	
		Neutral	4,38	1,865	247	
		Aufklärung	4,36	1,781	273	
		Negativ-Aktivierung	4,32	1,810	258	
Schlauchfiguren	männlich	Gesamt	4,41	1,815	1065	
		Aktivierung	6,07	1,079	119	
		Neutral	5,90	1,376	83	
		Aufklärung	5,68	1,570	108	
		Negativ-Aktivierung	5,89	1,581	85	
	weiblich	Gesamt	5,89	1,403	395	
		Aktivierung	5,04	1,786	168	
		Neutral	5,21	1,664	164	
		Aufklärung	5,24	1,649	165	
		Negativ-Aktivierung	5,17	1,833	173	
	Gesamt	Gesamt	5,16	1,734	670	
		Aktivierung	5,47	1,612	287	
			Neutral	5,45	1,604	247

	Aufklärung	5,41	1,629	273
	Negativ-Aktivierung	5,41	1,784	258
	Gesamt	5,43	1,656	1065

C) Selbstwirksamkeitserwartung (Meanscore)

Abhängige Variable:swe_msc					
Geschlecht	Stereotype Threat-Bedingungen	Risikoverhalten-Bedingungen	M	SD	N
männlich	Aktivierung	raten	3,03	,335	35
		ohne Instruktion	3,03	,357	50
		sicher	3,02	,458	36
		Gesamt	3,03	,381	121
	Neutral	raten	3,07	,482	35
		ohne Instruktion	3,06	,319	31
		sicher	3,06	,473	18
		Gesamt	3,07	,422	84
	Aufklärung	raten	2,93	,410	39
		ohne Instruktion	2,94	,307	29
		sicher	3,11	,428	41
		Gesamt	3,00	,399	109
	Negativ-Aktivierung	raten	2,90	,414	33
		ohne Instruktion	2,79	,447	18
		sicher	3,06	,424	33
		Gesamt	2,94	,433	84
	Gesamt	raten	2,98	,415	142
		ohne Instruktion	2,99	,359	128
		sicher	3,06	,439	128
		Gesamt	3,01	,407	398
weiblich	Aktivierung	raten	2,83	,401	58
		ohne Instruktion	2,88	,367	55
		sicher	2,94	,398	54
		Gesamt	2,88	,389	167
	Neutral	raten	2,88	,283	53
		ohne Instruktion	2,93	,433	56
		sicher	2,83	,435	59
		Gesamt	2,88	,393	168
	Aufklärung	raten	2,87	,444	64
		ohne Instruktion	2,84	,408	62

		sicher	2,83	,375	42
		Gesamt	2,85	,413	168
	Negativ-Aktivierung	raten	3,05	,470	50
		ohne Instruktion	2,80	,368	61
		sicher	2,86	,377	59
		Gesamt	2,89	,415	170
	Gesamt	raten	2,90	,413	225
		ohne Instruktion	2,86	,395	234
		sicher	2,86	,398	214
		Gesamt	2,88	,402	673
Gesamt	Aktivierung	raten	2,90	,389	93
		ohne Instruktion	2,95	,369	105
		sicher	2,97	,423	90
		Gesamt	2,94	,392	288
	Neutral	raten	2,96	,384	88
		ohne Instruktion	2,98	,400	87
		sicher	2,88	,452	77
		Gesamt	2,94	,411	252
	Aufklärung	raten	2,89	,430	103
		ohne Instruktion	2,88	,380	91
		sicher	2,97	,424	83
		Gesamt	2,91	,413	277
	Negativ-Aktivierung	raten	2,99	,452	83
		ohne Instruktion	2,79	,384	79
		sicher	2,93	,404	92
		Gesamt	2,91	,421	254
	Gesamt	raten	2,93	,415	367
		ohne Instruktion	2,90	,387	362
		sicher	2,94	,425	342
		Gesamt	2,93	,409	1071

D) Leistungen in EMS-ähnlichen Untertests

EMS-ähnl. Untertest	Geschlecht	Stereotype Threat-Bedingungen	Risikoverhalten-Bedingungen	M	SD	N
Quantitative u. formale Probleme	männlich	Aktivierung	raten	4,03	1,689	35
			ohne Instruktion	3,59	1,627	51
			sicher	2,70	1,714	37
			Gesamt	3,45	1,738	123
		Neutral	raten	3,77	1,767	35
			ohne Instruktion	3,19	1,621	31
			sicher	3,26	2,156	19
			Gesamt	3,45	1,809	85
		Aufklärung	raten	3,50	2,112	40
			ohne Instruktion	3,17	2,102	30
			sicher	3,24	1,574	42
			Gesamt	3,31	1,913	112
		Negativ-Aktivierung	raten	3,34	1,644	35
			ohne Instruktion	2,47	1,679	19
			sicher	3,29	1,601	35
			Gesamt	3,13	1,653	89
		Gesamt	raten	3,66	1,823	145
			ohne Instruktion	3,24	1,771	131
			sicher	3,11	1,711	133
			Gesamt	3,34	1,782	409
	weiblich	Aktivierung	raten	2,97	1,589	58
			ohne Instruktion	2,68	1,681	57
			sicher	2,85	1,638	55
			Gesamt	2,84	1,631	170
		Neutral	raten	3,31	1,632	55
			ohne Instruktion	2,73	1,635	56
			sicher	2,33	1,399	61
			Gesamt	2,77	1,597	172
		Aufklärung	raten	2,91	1,550	64
			ohne Instruktion	2,79	1,631	62
			sicher	2,88	1,435	42
			Gesamt	2,86	1,545	168
Negativ-Aktivierung		raten	2,75	1,246	51	
		ohne Instruktion	2,63	1,485	62	
		sicher	2,85	1,787	61	
		Gesamt	2,74	1,531	174	

Schlauchfiguren	Gesamt	Gesamt	raten	2,98	1,522	228
			ohne Instruktion	2,71	1,598	237
			sicher	2,71	1,590	219
			Gesamt	2,80	1,573	684
		Aktivierung	raten	3,37	1,699	93
			ohne Instruktion	3,11	1,709	108
			sicher	2,79	1,661	92
			Gesamt	3,09	1,701	293
		Neutral	raten	3,49	1,691	90
			ohne Instruktion	2,90	1,636	87
			sicher	2,55	1,645	80
			Gesamt	3,00	1,697	257
		Aufklärung	raten	3,13	1,801	104
			ohne Instruktion	2,91	1,795	92
			sicher	3,06	1,508	84
			Gesamt	3,04	1,713	280
	Negativ-Aktivierung	raten	2,99	1,443	86	
		ohne Instruktion	2,59	1,523	81	
		sicher	3,01	1,726	96	
		Gesamt	2,87	1,581	263	
	Gesamt	raten	3,24	1,676	373	
		ohne Instruktion	2,90	1,679	368	
		sicher	2,86	1,645	352	
		Gesamt	3,00	1,675	1093	
	männlich	Aktivierung	raten	12,11	3,066	35
			ohne Instruktion	11,86	2,892	51
			sicher	9,73	3,485	37
			Gesamt	11,29	3,271	123
Neutral		raten	11,40	3,274	35	
		ohne Instruktion	10,58	3,433	31	
		sicher	10,95	3,582	19	
		Gesamt	11,00	3,381	85	
Aufklärung		raten	12,35	2,806	40	
		ohne Instruktion	11,57	3,785	30	
		sicher	10,45	2,932	42	
		Gesamt	11,43	3,218	112	
Negativ-Aktivierung		raten	10,43	3,432	35	
		ohne Instruktion	11,00	3,249	19	
		sicher	11,60	3,882	35	

		Gesamt	11,01	3,579	89
	Gesamt	raten	11,60	3,198	145
		ohne Instruktion	11,37	3,296	131
		sicher	10,62	3,481	133
		Gesamt	11,21	3,341	409
weiblich		Aktivierung	raten	9,02	3,945
	ohne Instruktion		9,96	3,316	57
	sicher		10,13	3,389	55
	Gesamt		9,69	3,578	170
	Neutral	raten	9,15	3,768	55
		ohne Instruktion	9,14	3,534	56
		sicher	9,72	2,841	61
		Gesamt	9,35	3,378	172
	Aufklärung	raten	9,23	3,346	64
		ohne Instruktion	9,81	3,661	62
		sicher	10,71	2,831	42
		Gesamt	9,82	3,379	168
	Negativ-Aktivierung	raten	9,10	3,022	51
		ohne Instruktion	9,95	2,778	62
		sicher	9,33	3,520	61
		Gesamt	9,48	3,129	174
	Gesamt	raten	9,13	3,523	228
		ohne Instruktion	9,73	3,330	237
		sicher	9,90	3,198	219
		Gesamt	9,58	3,366	684
Gesamt	Aktivierung	raten	10,18	3,923	93
		ohne Instruktion	10,86	3,251	108
		sicher	9,97	3,414	92
		Gesamt	10,37	3,536	293
	Neutral	raten	10,02	3,733	90
		ohne Instruktion	9,66	3,546	87
		sicher	10,01	3,054	80
		Gesamt	9,89	3,461	257
	Aufklärung	raten	10,43	3,486	104
		ohne Instruktion	10,38	3,773	92
		sicher	10,58	2,867	84
		Gesamt	10,46	3,403	280
	Negativ-Aktivierung	raten	9,64	3,243	86
		ohne Instruktion	10,20	2,909	81

Med.nat. Grundv	männlich	Gesamt	sicher	10,16	3,798	96
			Gesamt	10,00	3,360	263
			raten	10,09	3,604	373
			ohne Instruktion	10,31	3,406	368
			sicher	10,18	3,321	352
		Gesamt	10,19	3,446	1093	
		Aktivierung	raten	3,26	1,837	35
			ohne Instruktion	3,12	1,423	51
			sicher	2,41	1,363	37
			Gesamt	2,94	1,564	123
	Neutral	raten	3,54	1,704	35	
		ohne Instruktion	2,68	1,904	31	
		sicher	3,00	2,028	19	
		Gesamt	3,11	1,871	85	
	Aufklärung	raten	3,47	1,648	40	
		ohne Instruktion	2,90	1,729	30	
		sicher	3,12	1,783	42	
		Gesamt	3,19	1,722	112	
	Negativ-Aktivierung	raten	3,09	1,772	35	
		ohne Instruktion	2,63	1,640	19	
sicher		2,66	1,494	35		
Gesamt		2,82	1,635	89		
Gesamt	raten	3,34	1,730	145		
	ohne Instruktion	2,89	1,642	131		
	sicher	2,78	1,648	133		
	Gesamt	3,02	1,690	409		
weiblich	Aktivierung	raten	2,64	1,398	58	
		ohne Instruktion	2,75	1,661	57	
		sicher	2,65	1,578	55	
		Gesamt	2,68	1,540	170	
	Neutral	raten	3,24	2,054	55	
		ohne Instruktion	2,48	1,768	56	
		sicher	2,62	1,624	61	
		Gesamt	2,77	1,836	172	
	Aufklärung	raten	2,73	1,828	64	
		ohne Instruktion	2,44	1,666	62	
		sicher	3,00	1,608	42	
		Gesamt	2,69	1,720	168	
	Negativ-Aktivierung	raten	2,63	1,523	51	

	Gesamt		ohne Instruktion	2,39	1,407	62	
			sicher	2,85	1,833	61	
			Gesamt	2,62	1,604	174	
		Gesamt	raten	2,81	1,730	228	
			ohne Instruktion	2,51	1,622	237	
			sicher	2,77	1,666	219	
		Gesamt	Gesamt	2,69	1,676	684	
			Aktivierung	raten	2,87	1,596	93
				ohne Instruktion	2,93	1,557	108
		sicher		2,55	1,493	92	
		Gesamt		2,79	1,553	293	
		Neutral	raten	3,36	1,922	90	
	ohne Instruktion		2,55	1,809	87		
	sicher		2,71	1,722	80		
	Gesamt		2,88	1,850	257		
	Aufklärung	raten	3,02	1,790	104		
		ohne Instruktion	2,59	1,691	92		
		sicher	3,06	1,689	84		
		Gesamt	2,89	1,735	280		
	Negativ-Aktivierung	raten	2,81	1,634	86		
		ohne Instruktion	2,44	1,458	81		
	Gesamt		sicher	2,78	1,712	96	
			Gesamt	2,69	1,614	263	
			Gesamt	raten	3,02	1,747	373
ohne Instruktion		2,65		1,637	368		
sicher		2,77		1,657	352		
Gesamt		2,81		1,688	1093		
männlich		Aktivierung	raten	3,80	1,922	35	
			ohne Instruktion	3,29	1,346	51	
			sicher	2,95	1,779	37	
			Gesamt	3,33	1,678	123	
		Neutral	raten	3,86	1,785	35	
			ohne Instruktion	3,06	1,692	31	
	sicher		3,11	1,853	19		
	Gesamt		3,40	1,788	85		
	Aufklärung	raten	3,08	1,347	40		
		ohne Instruktion	2,97	1,497	30		
		sicher	3,36	1,910	42		
		Gesamt	3,15	1,612	112		

		Negativ-Aktivierung	raten	3,34	1,589	35
			ohne Instruktion	2,58	1,742	19
			sicher	3,09	1,522	35
			Gesamt	3,08	1,604	89
		Gesamt	raten	3,50	1,680	145
			ohne Instruktion	3,06	1,528	131
			sicher	3,14	1,757	133
			Gesamt	3,24	1,666	409
weiblich	Aktivierung	raten	3,07	1,663	58	
		ohne Instruktion	2,81	1,705	57	
		sicher	2,95	1,580	55	
		Gesamt	2,94	1,645	170	
	Neutral	raten	3,33	1,689	55	
		ohne Instruktion	3,00	1,584	56	
		sicher	2,33	1,524	61	
		Gesamt	2,87	1,644	172	
	Aufklärung	raten	2,59	1,650	64	
		ohne Instruktion	3,29	1,653	62	
		sicher	2,62	1,513	42	
		Gesamt	2,86	1,643	168	
	Negativ-Aktivierung	raten	3,18	1,367	51	
		ohne Instruktion	2,47	1,197	62	
		sicher	2,84	1,714	61	
		Gesamt	2,80	1,465	174	
	Gesamt	raten	3,02	1,619	228	
		ohne Instruktion	2,89	1,564	237	
		sicher	2,68	1,599	219	
		Gesamt	2,87	1,597	684	
	Gesamt	Aktivierung	raten	3,34	1,791	93
			ohne Instruktion	3,04	1,558	108
			sicher	2,95	1,653	92
			Gesamt	3,11	1,667	293
Neutral		raten	3,53	1,737	90	
		ohne Instruktion	3,02	1,614	87	
		sicher	2,51	1,630	80	
		Gesamt	3,04	1,708	257	
Aufklärung		raten	2,78	1,551	104	
		ohne Instruktion	3,18	1,603	92	
		sicher	2,99	1,753	84	
		Gesamt	2,95	1,636	280	

Figuren lernen		Negativ-Aktivierung	Gesamt	2,98	1,634	280
			raten	3,24	1,455	86
			ohne Instruktion	2,49	1,333	81
			sicher	2,93	1,643	96
		Gesamt	2,90	1,516	263	
		Gesamt	raten	3,21	1,657	373
			ohne Instruktion	2,95	1,551	368
			sicher	2,85	1,673	352
	Gesamt		3,01	1,633	1093	
	männlich	Aktivierung	raten	4,63	1,942	35
			ohne Instruktion	4,29	1,879	51
			sicher	4,32	2,161	37
			Gesamt	4,40	1,974	123
		Neutral	raten	4,14	2,017	35
			ohne Instruktion	3,61	2,060	31
			sicher	4,58	2,269	19
			Gesamt	4,05	2,098	85
		Aufklärung	raten	4,68	1,639	40
			ohne Instruktion	4,17	2,547	30
			sicher	4,40	1,964	42
Gesamt			4,44	2,026	112	
Negativ-Aktivierung		raten	4,89	2,153	35	
		ohne Instruktion	4,05	2,272	19	
		sicher	4,37	2,438	35	
		Gesamt	4,51	2,292	89	
Gesamt		raten	4,59	1,935	145	
		ohne Instruktion	4,07	2,138	131	
		sicher	4,40	2,171	133	
		Gesamt	4,36	2,085	409	
weiblich	Aktivierung	raten	4,22	1,956	58	
		ohne Instruktion	4,40	2,137	57	
		sicher	4,49	1,961	55	
		Gesamt	4,37	2,011	170	
	Neutral	raten	4,49	2,227	55	
		ohne Instruktion	4,54	1,963	56	
		sicher	4,08	2,035	61	
		Gesamt	4,36	2,074	172	
	Aufklärung	raten	4,25	1,718	64	
		ohne Instruktion	5,39	2,043	62	

	Gesamt		sicher	4,17	1,793	42	
			Gesamt	4,65	1,937	168	
		Negativ-Aktivierung	raten	4,51	2,111	51	
			ohne Instruktion	4,81	1,845	62	
			sicher	4,20	2,337	61	
			Gesamt	4,51	2,109	174	
		Gesamt	raten	4,36	1,990	228	
			ohne Instruktion	4,80	2,021	237	
			sicher	4,23	2,056	219	
			Gesamt	4,47	2,034	684	
		Gesamt	Aktivierung	raten	4,38	1,950	93
				ohne Instruktion	4,35	2,011	108
				sicher	4,42	2,034	92
				Gesamt	4,38	1,992	293
			Neutral	raten	4,36	2,143	90
				ohne Instruktion	4,21	2,035	87
	sicher			4,20	2,089	80	
	Gesamt			4,26	2,083	257	
	Aufklärung		raten	4,41	1,693	104	
			ohne Instruktion	4,99	2,280	92	
			sicher	4,29	1,873	84	
			Gesamt	4,56	1,972	280	
	Negativ-Aktivierung		raten	4,66	2,123	86	
			ohne Instruktion	4,63	1,965	81	
			sicher	4,26	2,363	96	
			Gesamt	4,51	2,168	263	
	Gesamt		raten	4,45	1,969	373	
			ohne Instruktion	4,54	2,090	368	
sicher			4,30	2,098	352		
Gesamt			4,43	2,053	1093		
Fakten lernen	männlich	Aktivierung	raten	4,43	1,836	35	
			ohne Instruktion	4,14	1,767	51	
			sicher	3,51	1,895	37	
			Gesamt	4,03	1,846	123	
		Neutral	raten	3,63	1,573	35	
			ohne Instruktion	3,77	2,186	31	
			sicher	3,47	1,837	19	
	Gesamt		3,65	1,856	85		
	Aufklärung	raten	4,38	1,764	40		

			ohne Instruktion	4,23	1,591	30
			sicher	3,81	1,534	42
			Gesamt	4,12	1,639	112
		Negativ-Aktivierung	raten	3,97	1,339	35
			ohne Instruktion	3,89	1,410	19
			sicher	3,77	2,102	35
			Gesamt	3,88	1,678	89
		Gesamt	raten	4,11	1,659	145
			ohne Instruktion	4,04	1,782	131
			sicher	3,67	1,825	133
			Gesamt	3,94	1,760	409
	weiblich	Aktivierung	raten	4,53	1,536	58
			ohne Instruktion	4,26	2,109	57
			sicher	4,62	1,683	55
			Gesamt	4,47	1,788	170
		Neutral	raten	4,36	1,778	55
			ohne Instruktion	4,14	1,678	56
			sicher	4,30	1,865	61
			Gesamt	4,27	1,770	172
		Aufklärung	raten	4,39	1,443	64
			ohne Instruktion	4,69	1,843	62
			sicher	4,62	1,710	42
			Gesamt	4,56	1,662	168
		Negativ-Aktivierung	raten	4,47	1,859	51
			ohne Instruktion	4,65	1,631	62
			sicher	4,05	1,802	61
			Gesamt	4,39	1,769	174
		Gesamt	raten	4,44	1,639	228
			ohne Instruktion	4,45	1,826	237
			sicher	4,37	1,778	219
			Gesamt	4,42	1,748	684
	Gesamt	Aktivierung	raten	4,49	1,646	93
			ohne Instruktion	4,20	1,947	108
			sicher	4,17	1,843	92
			Gesamt	4,29	1,823	293
		Neutral	raten	4,08	1,730	90
			ohne Instruktion	4,01	1,871	87
			sicher	4,10	1,880	80
			Gesamt	4,06	1,819	257

		Aufklärung	raten	4,38	1,566	104
			ohne Instruktion	4,54	1,769	92
			sicher	4,21	1,665	84
			Gesamt	4,39	1,664	280
		Negativ-Aktivierung	raten	4,27	1,676	86
			ohne Instruktion	4,47	1,605	81
			sicher	3,95	1,910	96
			Gesamt	4,21	1,752	263
		Gesamt	raten	4,31	1,653	373
			ohne Instruktion	4,30	1,818	368
			sicher	4,11	1,825	352
			Gesamt	4,24	1,767	1093

E) Leistungen in EMS-ähnlichen Untertests unter Berücksichtigung der SWE als Kovariate

EMS-ähnl. Untertests	Geschlecht	Stereotype Threat-Bedingungen	Risikoverhalten-Bedingungen	M	SD	N
Quantitative u. formale Probleme	männlich	Aktivierung	raten	4,03	1,689	35
			ohne Instruktion	3,62	1,627	50
			sicher	2,78	1,675	36
			Gesamt	3,49	1,718	121
		Neutral	raten	3,77	1,767	35
			ohne Instruktion	3,19	1,621	31
			sicher	3,33	2,196	18
			Gesamt	3,46	1,813	84
		Aufklärung	raten	3,56	2,100	39
			ohne Instruktion	3,17	2,139	29
			sicher	3,32	1,507	41
			Gesamt	3,37	1,899	109
		Negativ-Aktivierung	raten	3,45	1,622	33
			ohne Instruktion	2,50	1,724	18
			sicher	3,21	1,576	33
			Gesamt	3,15	1,646	84
		Gesamt	raten	3,70	1,809	142
			ohne Instruktion	3,26	1,785	128
			sicher	3,14	1,678	128
			Gesamt	3,38	1,773	398

weiblich	Aktivierung	raten	2,97	1,589	58
		ohne Instruktion	2,71	1,696	55
		sicher	2,89	1,633	54
		Gesamt	2,86	1,633	167
	Neutral	raten	3,40	1,585	53
		ohne Instruktion	2,73	1,635	56
		sicher	2,34	1,409	59
		Gesamt	2,80	1,594	168
	Aufklärung	raten	2,91	1,550	64
		ohne Instruktion	2,79	1,631	62
		sicher	2,88	1,435	42
		Gesamt	2,86	1,545	168
	Negativ-Aktivierung	raten	2,78	1,234	50
		ohne Instruktion	2,61	1,486	61
		sicher	2,90	1,797	59
		Gesamt	2,76	1,533	170
	Gesamt	raten	3,01	1,512	225
		ohne Instruktion	2,71	1,602	234
		sicher	2,74	1,594	214
		Gesamt	2,82	1,574	673
Gesamt	Aktivierung	raten	3,37	1,699	93
		ohne Instruktion	3,14	1,718	105
		sicher	2,84	1,642	90
		Gesamt	3,12	1,695	288
	Neutral	raten	3,55	1,660	88
		ohne Instruktion	2,90	1,636	87
		sicher	2,57	1,666	77
		Gesamt	3,02	1,696	252
	Aufklärung	raten	3,16	1,797	103
		ohne Instruktion	2,91	1,805	91
		sicher	3,10	1,478	83
		Gesamt	3,06	1,708	277
	Negativ-Aktivierung	raten	3,05	1,431	83
		ohne Instruktion	2,58	1,533	79
		sicher	3,01	1,719	92
		Gesamt	2,89	1,579	254
	Gesamt	raten	3,28	1,666	367
		ohne Instruktion	2,90	1,687	362
		sicher	2,89	1,635	342

Schlauchfiguren	männlich	Aktivierung	Gesamt	3,03	1,672	1071
			raten	12,11	3,066	35
			ohne Instruktion	11,98	2,796	50
			sicher	9,94	3,277	36
		Gesamt	11,41	3,148	121	
		Neutral	raten	11,40	3,274	35
			ohne Instruktion	10,58	3,433	31
			sicher	11,00	3,678	18
			Gesamt	11,01	3,399	84
		Aufklärung	raten	12,38	2,834	39
			ohne Instruktion	11,69	3,790	29
			sicher	10,46	2,967	41
			Gesamt	11,48	3,242	109
		Negativ-Aktivierung	raten	10,36	3,507	33
			ohne Instruktion	10,94	3,334	18
			sicher	11,52	3,985	33
			Gesamt	10,94	3,661	84
		Gesamt	raten	11,61	3,227	142
			ohne Instruktion	11,43	3,286	128
			sicher	10,66	3,451	128
	Gesamt		11,25	3,336	398	
	weiblich	Aktivierung	raten	9,02	3,945	58
			ohne Instruktion	9,98	3,364	55
			sicher	10,15	3,417	54
			Gesamt	9,70	3,606	167
		Neutral	raten	9,23	3,630	53
			ohne Instruktion	9,14	3,534	56
			sicher	9,66	2,832	59
			Gesamt	9,35	3,325	168
		Aufklärung	raten	9,23	3,346	64
			ohne Instruktion	9,81	3,661	62
			sicher	10,71	2,831	42
Gesamt			9,82	3,379	168	
Negativ-Aktivierung		raten	9,06	3,040	50	
		ohne Instruktion	10,00	2,775	61	
		sicher	9,46	3,405	59	
		Gesamt	9,54	3,089	170	
Gesamt		raten	9,14	3,493	225	
		ohne Instruktion	9,74	3,344	234	

Med. nat. Grundv.	Gesamt		sicher	9,93	3,163	214
			Gesamt	9,60	3,351	673
		Aktivierung	raten	10,18	3,923	93
			ohne Instruktion	10,93	3,250	105
			sicher	10,07	3,345	90
			Gesamt	10,42	3,519	288
		Neutral	raten	10,09	3,635	88
			ohne Instruktion	9,66	3,546	87
			sicher	9,97	3,078	77
			Gesamt	9,90	3,434	252
		Aufklärung	raten	10,43	3,502	103
			ohne Instruktion	10,41	3,786	91
			sicher	10,59	2,884	83
			Gesamt	10,47	3,418	277
		Negativ-Aktivierung	raten	9,58	3,276	83
			ohne Instruktion	10,22	2,916	79
			sicher	10,20	3,736	92
			Gesamt	10,00	3,348	254
		Gesamt	raten	10,09	3,595	367
			ohne Instruktion	10,34	3,416	362
	sicher		10,21	3,288	342	
	Gesamt		10,21	3,437	1071	
	männlich	Aktivierung	raten	3,26	1,837	35
			ohne Instruktion	3,10	1,432	50
			sicher	2,47	1,320	36
			Gesamt	2,96	1,551	121
		Neutral	raten	3,54	1,704	35
			ohne Instruktion	2,68	1,904	31
sicher			2,94	2,071	18	
Gesamt			3,10	1,880	84	
Aufklärung		raten	3,56	1,569	39	
		ohne Instruktion	2,97	1,721	29	
		sicher	3,17	1,773	41	
		Gesamt	3,26	1,691	109	
Negativ-Aktivierung		raten	3,09	1,826	33	
		ohne Instruktion	2,67	1,680	18	
		sicher	2,73	1,506	33	
		Gesamt	2,86	1,666	84	
Gesamt		raten	3,37	1,724	142	

			ohne Instruktion	2,91	1,648	128
			sicher	2,83	1,642	128
			Gesamt	3,05	1,687	398
	weiblich	Aktivierung	raten	2,64	1,398	58
			ohne Instruktion	2,80	1,649	55
			sicher	2,69	1,576	54
			Gesamt	2,71	1,534	167
		Neutral	raten	3,36	1,991	53
			ohne Instruktion	2,48	1,768	56
			sicher	2,61	1,651	59
			Gesamt	2,80	1,832	168
		Aufklärung	raten	2,73	1,828	64
			ohne Instruktion	2,44	1,666	62
			sicher	3,00	1,608	42
			Gesamt	2,69	1,720	168
		Negativ-Aktivierung	raten	2,64	1,535	50
			ohne Instruktion	2,41	1,407	61
			sicher	2,92	1,822	59
			Gesamt	2,65	1,603	170
		Gesamt	raten	2,84	1,720	225
			ohne Instruktion	2,53	1,621	234
			sicher	2,79	1,669	214
			Gesamt	2,71	1,673	673
	Gesamt	Aktivierung	raten	2,87	1,596	93
			ohne Instruktion	2,94	1,549	105
			sicher	2,60	1,475	90
			Gesamt	2,81	1,544	288
		Neutral	raten	3,43	1,874	88
			ohne Instruktion	2,55	1,809	87
			sicher	2,69	1,749	77
			Gesamt	2,90	1,849	252
		Aufklärung	raten	3,05	1,773	103
			ohne Instruktion	2,60	1,692	91
			sicher	3,08	1,684	83
			Gesamt	2,91	1,728	277
		Negativ-Aktivierung	raten	2,82	1,661	83
			ohne Instruktion	2,47	1,466	79
			sicher	2,85	1,709	92
			Gesamt	2,72	1,624	254

Diagramme u. Tabellen		Gesamt	raten	3,04	1,739	367
			ohne Instruktion	2,66	1,639	362
			sicher	2,80	1,657	342
			Gesamt	2,84	1,685	1071
	männlich	Aktivierung	raten	3,80	1,922	35
			ohne Instruktion	3,32	1,347	50
			sicher	3,03	1,732	36
			Gesamt	3,37	1,659	121
		Neutral	raten	3,86	1,785	35
			ohne Instruktion	3,06	1,692	31
			sicher	3,11	1,906	18
			Gesamt	3,40	1,798	84
		Aufklärung	raten	3,15	1,268	39
			ohne Instruktion	3,07	1,412	29
			sicher	3,37	1,933	41
			Gesamt	3,21	1,576	109
		Negativ-Aktivierung	raten	3,39	1,619	33
			ohne Instruktion	2,56	1,790	18
			sicher	3,27	1,353	33
			Gesamt	3,17	1,574	84
		Gesamt	raten	3,54	1,666	142
			ohne Instruktion	3,09	1,519	128
			sicher	3,21	1,724	128
			Gesamt	3,29	1,646	398
	weiblich	Aktivierung	raten	3,07	1,663	58
			ohne Instruktion	2,87	1,689	55
			sicher	2,96	1,590	54
			Gesamt	2,97	1,641	167
Neutral		raten	3,45	1,588	53	
		ohne Instruktion	3,00	1,584	56	
		sicher	2,31	1,545	59	
		Gesamt	2,90	1,633	168	
Aufklärung		raten	2,59	1,650	64	
		ohne Instruktion	3,29	1,653	62	
		sicher	2,62	1,513	42	
		Gesamt	2,86	1,643	168	
Negativ-Aktivierung		raten	3,18	1,380	50	
		ohne Instruktion	2,48	1,206	61	
		sicher	2,92	1,684	59	

	Gesamt	Gesamt	Gesamt	2,84	1,458	170
			raten	3,05	1,604	225
			ohne Instruktion	2,91	1,560	234
			sicher	2,70	1,602	214
			Gesamt	2,89	1,592	673
	Gesamt	Aktivierung	raten	3,34	1,791	93
			ohne Instruktion	3,09	1,545	105
			sicher	2,99	1,639	90
			Gesamt	3,14	1,657	288
		Neutral	raten	3,61	1,671	88
			ohne Instruktion	3,02	1,614	87
			sicher	2,49	1,659	77
			Gesamt	3,07	1,703	252
		Aufklärung	raten	2,81	1,534	103
			ohne Instruktion	3,22	1,576	91
			sicher	2,99	1,763	83
			Gesamt	3,00	1,623	277
		Negativ-Aktivierung	raten	3,27	1,474	83
			ohne Instruktion	2,49	1,348	79
			sicher	3,04	1,575	92
Gesamt			2,94	1,503	254	
Gesamt		raten	3,24	1,644	367	
		ohne Instruktion	2,98	1,546	362	
		sicher	2,89	1,665	342	
		Gesamt	3,04	1,624	1071	
Figuren lernen	männlich	Aktivierung	raten	4,63	1,942	35
			ohne Instruktion	4,28	1,896	50
			sicher	4,36	2,180	36
			Gesamt	4,40	1,986	121
		Neutral	raten	4,14	2,017	35
			ohne Instruktion	3,61	2,060	31
			sicher	4,67	2,301	18
			Gesamt	4,06	2,108	84
		Aufklärung	raten	4,64	1,646	39
			ohne Instruktion	4,07	2,534	29
			sicher	4,39	1,986	41
			Gesamt	4,39	2,032	109
		Negativ-Aktivierung	raten	4,91	2,185	33
			ohne Instruktion	3,83	2,121	18

		sicher	4,64	2,247	33	
		Gesamt	4,57	2,208	84	
	Gesamt	raten	4,58	1,944	142	
		ohne Instruktion	4,01	2,117	128	
		sicher	4,48	2,133	128	
		Gesamt	4,36	2,072	398	
weiblich	Aktivierung	raten	4,22	1,956	58	
		ohne Instruktion	4,47	2,089	55	
		sicher	4,46	1,969	54	
		Gesamt	4,38	1,996	167	
	Neutral	raten	4,60	2,169	53	
		ohne Instruktion	4,54	1,963	56	
		sicher	4,03	2,051	59	
		Gesamt	4,38	2,064	168	
	Aufklärung	raten	4,25	1,718	64	
		ohne Instruktion	5,39	2,043	62	
		sicher	4,17	1,793	42	
		Gesamt	4,65	1,937	168	
	Negativ-Aktivierung	raten	4,50	2,131	50	
		ohne Instruktion	4,77	1,838	61	
		sicher	4,34	2,241	59	
		Gesamt	4,54	2,067	170	
	Gesamt	raten	4,38	1,979	225	
		ohne Instruktion	4,81	2,005	234	
		sicher	4,25	2,031	214	
		Gesamt	4,49	2,016	673	
	Gesamt	Aktivierung	raten	4,38	1,950	93
			ohne Instruktion	4,38	1,992	105
			sicher	4,42	2,044	90
			Gesamt	4,39	1,988	288
Neutral		raten	4,42	2,110	88	
		ohne Instruktion	4,21	2,035	87	
		sicher	4,18	2,113	77	
		Gesamt	4,27	2,080	252	
Aufklärung		raten	4,40	1,694	103	
		ohne Instruktion	4,97	2,283	91	
		sicher	4,28	1,883	83	
		Gesamt	4,55	1,975	277	
Negativ-Aktivierung		raten	4,66	2,149	83	

Fakten lernen			ohne Instruktion	4,56	1,933	79	
			sicher	4,45	2,235	92	
			Gesamt	4,55	2,110	254	
		Gesamt	raten	4,46	1,965	367	
			ohne Instruktion	4,52	2,078	362	
			sicher	4,34	2,070	342	
			Gesamt	4,44	2,037	1071	
		männlich	Aktivierung	raten	4,43	1,836	35
				ohne Instruktion	4,14	1,784	50
				sicher	3,61	1,825	36
				Gesamt	4,07	1,825	121
			Neutral	raten	3,63	1,573	35
				ohne Instruktion	3,77	2,186	31
				sicher	3,39	1,852	18
				Gesamt	3,63	1,861	84
			Aufklärung	raten	4,46	1,699	39
				ohne Instruktion	4,10	1,448	29
				sicher	3,80	1,553	41
				Gesamt	4,12	1,591	109
			Negativ-Aktivierung	raten	4,03	1,357	33
			ohne Instruktion	3,89	1,451	18	
			sicher	4,00	1,936	33	
			Gesamt	3,99	1,609	84	
		Gesamt	raten	4,15	1,650	142	
			ohne Instruktion	4,01	1,768	128	
			sicher	3,74	1,767	128	
			Gesamt	3,97	1,730	398	
	weiblich	Aktivierung	raten	4,53	1,536	58	
			ohne Instruktion	4,33	2,064	55	
			sicher	4,70	1,574	54	
			Gesamt	4,52	1,735	167	
		Neutral	raten	4,47	1,694	53	
			ohne Instruktion	4,14	1,678	56	
			sicher	4,20	1,827	59	
			Gesamt	4,27	1,732	168	
		Aufklärung	raten	4,39	1,443	64	
			ohne Instruktion	4,69	1,843	62	
			sicher	4,62	1,710	42	
			Gesamt	4,56	1,662	168	

		Negativ-Aktivierung	raten	4,50	1,865	50
			ohne Instruktion	4,64	1,644	61
			sicher	4,19	1,666	59
			Gesamt	4,44	1,720	170
		Gesamt	raten	4,47	1,618	225
			ohne Instruktion	4,46	1,813	234
			sicher	4,41	1,703	214
			Gesamt	4,45	1,712	673
Gesamt	Aktivierung	Aktivierung	raten	4,49	1,646	93
			ohne Instruktion	4,24	1,929	105
			sicher	4,27	1,753	90
			Gesamt	4,33	1,784	288
	Neutral	Neutral	raten	4,14	1,690	88
			ohne Instruktion	4,01	1,871	87
			sicher	4,01	1,853	77
			Gesamt	4,06	1,798	252
	Aufklärung	Aufklärung	raten	4,42	1,537	103
			ohne Instruktion	4,51	1,741	91
			sicher	4,22	1,675	83
			Gesamt	4,39	1,646	277
	Negativ-Aktivierung	Negativ-Aktivierung	raten	4,31	1,689	83
			ohne Instruktion	4,47	1,624	79
			sicher	4,12	1,759	92
			Gesamt	4,29	1,694	254
	Gesamt	Gesamt	raten	4,35	1,636	367
			ohne Instruktion	4,30	1,808	362
			sicher	4,16	1,754	342
			Gesamt	4,27	1,734	1071

Lebenslauf

Persönliche Daten

Name: Anita Teufl
Geburtsort: Amstetten

Schulbildung

1999–2004 Höhere Bundeslehranstalt für wirtschaftliche Berufe
Amstetten
1995–1999 Hauptschule Euratsfeld
1991–1995 Volksschule Ferschnitz

Studium

Studienjahr 2008/09 Leistungsstipendium der Universität Wien
seit 03/2007 zweiter Studienabschnitt, Wahlpflichtfächer:
Klinische- und Gesundheitspsychologie, Angewandte
Kinder- und Jugendpsychologie
Freies Wahlfach: Absolvierung an der Medizinischen
Universität Wien mit dem Schwerpunkt Psychiatrie
02/2007 Abschluss des ersten Studienabschnittes
Studienjahr 2004/05 Leistungsstipendium der Universität Wien
seit 10/2004 Studium der Psychologie, Universität Wien

Berufserfahrung / Praktika

01/2010 – 07/2013 Universitätsklinik für Neurologie, Allgemeines
Krankenhaus der Stadt Wien
wissenschaftliche Mitarbeit
10/2009 – 09/2011 Krisenzentrum Amstetten
Aushilfspersonal, stationäre Betreuung von Jugendlichen
im Turnusdienst

Sommer 2011; 2010	Buntmetall Amstetten Ferialarbeit, Schichtarbeit im Produktionswerk
08/2009 – 09/2009	Krisenzentrum Amstetten psychologisches Praktikum
Sommer 2008; 2007; 2006; 2005; 2004	Buntmetall Amstetten Ferialarbeit, Schichtarbeit im Produktionswerk
07/2002 – 09/2002	Amadeus Buchhandlung (Thalia), Amstetten Büropraktikum
08/2001	Buntmetall Amstetten Büropraktikum
07/2000	Lettershop Bruckner, Steinakirchen Praktikum

Weitere Kenntnisse

Fremdsprachen	Englisch und Französisch
EDV-Anwenderkenntnisse	PASW (SPSS), Word, Excel, PowerPoint, Access ECDL (Europäischer Computer Führerschein)
Führerschein	B
Interessen	Natur, Malen, Reisen, Basketball, Wandern