



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Arbeit/ Title of Diploma Thesis

Der Einfluss von Stereotype Threat auf die mathematischen Leistungen von Schülern und Schülerinnen der 9. Schulstufe

Verfasserin / Submitted by
Janine Martens

Angestrebter akademischer Grad / in partial fulfilment of the requirements for the de-
gree of

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2015 / Vienna, 2015

Studienkennzahl /
degree program code as it appears
on the student record sheet::
Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet::
Betreuer / Supervisor:

A 298

Psychologie

Ass.-Prof. Dr. Marco Jirasko

Danksagung

Ich möchte mich bei Herrn Ass.-Prof. Dr. Marco Jirasko für die fachliche Unterstützung und Leitung, sowie den Direktoren, Lehrern und Lehrerinnen sowie Schülern und Schülerinnen der Gymnasien für ihr persönliches Engagement und ihre Hilfsbereitschaft bei der Durchführung des Experiments bedanken. Dank gebührt auch meiner Studienkollegin Regina Reischer für Ihre Hilfe bei der gemeinsamen Erarbeitung des Testverfahrens sowie Ihre Unterstützung bei den gemeinsam durchgeführten Testungen.

INHALT

EINLEITUNG	1
1 THEORETISCHER TEIL	5
1.1 BEGRIFFSDEFINITION PRIMING.....	5
1.2 STEREOTYPE: BEGRIFFSDEFINITION, URSPRUNG UND FUNKTIONEN.....	7
1.2.1 <i>Begriffsdefinition Stereotyp</i>	8
1.2.2 <i>Entstehung von Stereotypen</i>	9
1.2.3 <i>Funktionen von Stereotypen</i>	9
1.3 GESCHLECHTSSTEREOTYPE UND GESCHLECHTERROLLEN	11
1.4 DAS DISSOZIATIVE MODELL NACH DEVINE	15
1.5 GESCHLECHTSUNTERSCHIEDE BEI MATHEMATISCHEN FÄHIGKEITEN	18
1.6 STEREOTYPE THREAT	20
1.6.1 <i>Stereotype Threat im Zusammenhang mit Priming</i>	22
1.6.2 <i>Moderatoren für Stereotype Threat</i>	24
1.6.3 <i>Integriertes Prozessmodel der Stereotype Threat Effekte auf die Leistung</i>	29
1.6.4 <i>Mediatoren von Stereotype Threat</i>	31
1.7 STEREOTYPE THREAT IM BEZUG AUF GENDER UND MATHEMATIKLEISTUNG.	34
1.8 STEREOTYPE LIFT	35
1.9 HYPOTHESEN.....	38
1.9.1 <i>H1: Gibt es einen signifikanten Unterschied in der Basisfähigkeit Mathematik zwischen Buben und Mädchen?</i>	38
1.9.2 <i>H2: Weibliche Versuchspersonen, welche einen Stereotypenkonformen (SL) Artikel erhalten, erbringen bessere oder</i>	

<i>gleich gute Leistungen bei dem Mathematiktest nach dem Priming, wie im Mathematiktest vor dem Priming</i>	<i>39</i>
<i>1.9.3 H3: Weibliche Versuchsteilnehmer, welche einen Stereotypkonformen Artikel (Stereotype Threat) erhalten, bringen im Mathematiktest nach dem Priming schlechtere Leistungen als im Mathematiktest vor dem Priming</i>	<i>39</i>
<i>1.9.4 H4: Weibliche Teilnehmer schneiden im Mathematiktest nach dem Priming unter Stereotype Lift besser ab als unter Stereotype Threat</i>	<i>40</i>
<i>1.9.5 H5: Männliche Versuchsteilnehmer, welche der Stereotype Lift Bedingung ausgesetzt werden, bringen genauso gute Leistungen im Mathematiktest vor, wie im Mathematiktest nach dem Priming</i>	<i>40</i>
<i>1.9.6 H6: Männliche Versuchsteilnehmer, welche der Stereotype Threat Bedingung ausgesetzt werden, bringen im Mathematiktest nach dem Priming bessere Leistungen, als im Mathematiktest vor dem Priming</i>	<i>40</i>
<i>1.9.7 H7: Männliche Versuchsteilnehmer in der Stereotype Threat Bedingung schneiden im Mathematiktest nach dem Priming gleich gut ab wie männliche Versuchsteilnehmer in der Stereotype Lift Bedingung.</i>	<i>41</i>
<i>1.9.8 H8: Männliche Versuchsteilnehmer schneiden im Mathematiktest nach dem Priming unter ST besser ab als Frauen unter ST</i>	<i>41</i>
<i>1.10 WEITERE FRAGESTELLUNGEN</i>	<i>41</i>
<i>1.10.1 Hat Interesse an Mathematik einen Einfluss auf Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde?</i>	<i>41</i>

1.10.2	<i>Hat die persönliche Wichtigkeit von Mathematik einen Effekt auf die Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde?</i>	42
1.10.3	<i>Beeinflusst das Priming die Selbsteinschätzung?</i>	42
1.10.4	<i>Beeinflusst das Priming die Fremdeinschätzung?</i>	42
1.10.5	<i>Wenn Teilnehmer eine sehr geringe Motivation aufweisen, hat das Priming keinen Einfluss auf die Mathematiktestwerte der Teilnehmer</i>	42
1.10.6	<i>Hat die Primingmethode einen Einfluss auf die Stärke des Stereotype Threat, oder Stereotype Lift Effekt?</i>	43
2	METHODE	44
2.1	UNTERSUCHUNGSDESIGN	44
2.1.1	<i>Durchführung der Untersuchung</i>	44
2.2	UNTERSUCHUNGSTRUMENTE	45
2.2.1	<i>CFT 20 R Subtest Matrizen</i>	45
2.2.2	<i>Baseline Mathematiktest und Parallel Versionen</i>	46
2.2.3	<i>Priming mittels Artikel</i>	48
2.2.4	<i>Fragebogen zum Arbeitsstil</i>	49
2.3	VORGEHEN	50
2.3.1	<i>Ablauf des ersten Testtermins</i>	50
2.3.2	<i>Ablauf des zweiten Testzeitpunktes</i>	54
2.3.3	<i>LimeSurvey</i>	57
2.3.4	<i>Der persönliche Code</i>	57
2.4	BESCHREIBUNG DER STICHPROBE	58
2.5	ANALYSE DER INSTRUMENTE	59

2.5.1	<i>CFT 20 R Subtest Matrizen</i>	59
2.5.2	<i>Baseline Mathematiktest – Pre Priming</i>	61
2.5.3	<i>Mathematiktest – Post-Priming</i>	63
2.5.4	<i>Kurzfragebogen zu den Arbeitshaltungen</i>	66
3	ERGEBNISSE	68
3.1	HAUPTHYPOTHESEN	68
3.1.1	<i>H1: Gibt es einen signifikanten Unterschied in der Basisfertigkeit Mathematik zwischen Mädchen und Burschen?</i>	68
3.1.2	<i>CFT – Unterschiede in der Fähigkeit des logischen Denkens Teilnehmer</i>	70
3.1.3	<i>Haupteffekt des Primings</i>	70
3.1.4	<i>H2: Weibliche Versuchspersonen, welche den stereotypnonkonformen Artikel (SL) bekommen, erbringen bessere oder gleich gute Leistungen bei dem Mathematiktest nach dem Priming wie im Mathematiktest vor dem Priming</i>	74
3.1.5	<i>H3: Weibliche Versuchsteilnehmer welche einen stereotypkonformen Artikel (Stereotype Threat) erhalten, bringen im Mathematiktest nach dem Priming schlechtere Leistungen als im Mathematiktest vor dem Priming</i>	76
3.1.6	<i>H4: Weibliche Teilnehmer schneiden im Mathematiktest nach dem Priming unter Stereotype Lift besser ab, als unter Stereotype Threat</i>	78
3.1.7	<i>H5: Männliche Versuchsteilnehmer, welche der Stereotype Lift Bedingung ausgesetzt werden, bringen genauso gute Leistungen im Mathematiktest vor, wie im Mathematiktest nach dem Priming</i>	78

3.1.8	<i>H6: Männliche Versuchsteilnehmer welche der Stereotype Threat Bedingung ausgesetzt werden, bringen im Mathematiktest nach dem Priming bessere Leistungen als im Mathematiktest vor dem Priming.....</i>	81
3.1.9	<i>H7: Männliche Versuchsteilnehmer in der Stereotype Threat Bedingung schneiden im Mathematiktest nach dem Priming, gleich gut ab wie männliche Versuchsteilnehmer in der Stereotype Lift Bedingung.</i>	83
3.1.10	<i>H8: Männliche Versuchsteilnehmer schneiden unter ST besser ab als Frauen unter ST.</i>	84
3.2	WEITERE FRAGESTELLUNGEN	85
3.2.1	<i>Hat Interesse an Mathematik einen Einfluss auf Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde</i>	85
3.2.2	<i>Hat die persönliche Wichtigkeit von Mathematik einen Effekt auf die Leistungen im Mathematiktest welcher vor dem Priming bearbeitet wurde?</i>	85
3.2.3	<i>Beeinflusst das Priming die Selbsteinschätzung?.....</i>	86
3.2.4	<i>Beeinflusst das Priming die Fremdeinschätzung?</i>	89
3.2.5	<i>Wenn die Teilnehmer eine geringe Motivation haben, hat das Priming keinen Einfluss auf die Mathematiktestwerte der Teilnehmer</i>	93
3.2.6	<i>Hat die Primingmethode einen Einfluss auf die Stärke des Stereotype Threat, oder Stereotype Lift Effekt?.....</i>	94
3.3	WEITERE ERGEBNISSE	94
3.3.1	<i>Stereotype Einstellungen bezüglich Mathematik vor der Bearbeitung von Aufgaben</i>	94
3.3.2	<i>Geschlechtsunterschiede bezüglich der Arbeitshaltungen.....</i>	95

3.3.3 Zusammenhänge zwischen den Arbeitshaltungen	97
4 DISKUSSION	99
4.1 ERGEBNISSE DER UNTERSUCHUNG	99
4.2 KRITIK UND AUSBLICK.....	101
LITERATUR	105
TABELLENVERZEICHNIS	114
ABBILDUNGSVERZEICHNIS	119
ANHANG.....	121
ANHANG A – ABSTRACT	121
ANHANG B – ZUSAMMENFASSUNG.....	122
ANHANG C – UNTERSUCHUNGSINSTRUMENTE	123
ANHANG D – HISTOGRAMME.....	140
ANHANG E – TABELLEN	143
ANHANG F – LEBENSLAUF	146

Einleitung

Viele Vorurteile ziehen sich durch verschiedene Gesellschaftsschichten und halten sich manches Mal über Generationen hinweg. Diese Vorurteile können sowohl positiver, als auch negativer Natur sein. Bergmann (2001) definierte den Begriff Vorurteil wie folgt:

„Im Alltagsverständnis gebrauchen wir den Begriff Vorurteil, um ausgeprägte positive und negative Urteile oder Einstellungen eines Mitmenschen über ein Vorurteilsobjekt zu bezeichnen, wenn wir sie für nicht realitätsgerecht halten und der Betreffende trotz Gegenargumenten nicht von seiner Meinung abrückt. Da wir in unseren Urteilen zumeist nur unsere Sichtweise wiedergeben und Urteile fast immer gewisse Verallgemeinerungen enthalten, sind in jedem Urteil Momente des Vorurteilshaften zu finden.“ (Bergmann, 2001, S. 1)

Man kann Vorurteile als Einstellungen, deren Objekte Außengruppen sind, sehen. In den meisten Fällen handelt es sich hierbei um negative Einstellungen. „Die kognitive Komponente der Vorurteile – das subjektive Wissen bzw. die Meinungen über die Außengruppe – wird Stereotyp genannt“ (Herkner, 2003, S. 493). Diese Vorurteile können durch bestimmte Reize auch gezielt ins Bewusstsein gerufen werden. Durch Voraktivierung eines Gedächtnisinhalts mittels eines Hinweisreizes, können bestimmte Assoziationen geweckt werden. Diesen Prozess nennt man *Priming* (Herkner, 2003).

In einem Experiment zeigten Bargh, Chen, und Burrows, (1996), dass Menschen die auf das Thema „Altern“ geprimt wurden, sich langsamer einen Flur entlang bewegten, als jene die mit neutralen Reizen konfrontiert worden waren. Geprimt wurde die Versuchsgruppe durch die Beschäftigung mit Worten, die mit dem Thema „Altern“ besetzt sind. Dies geschah unter dem Vorwand eines „Scrambled sentence Test“.

Es gilt zu erwähnen, dass es durchaus auch Vorurteile mit einem positiven Inhalt wie „Asiaten sind besonders gut in Mathe“ gibt. In Studien konnte nachgewiesen werden, dass auch diese positiven Stereotype einen Einfluss auf das Verhalten haben können. So konnten Aronson, Lustina, Good, Keough, Steele und Brown (1999) in einem Experiment nachweisen, dass weiße Männer, welche die Information erhielten, dass Asiaten in mathematischen Bereichen besser abschneiden, in einem anschließenden Test signifikant schlechter abschnitten, als eine Kontrollgruppe. In diesem Zusammenhang wird von „Stereotype Threat“ gesprochen. Besonders verbreitete Vorurteile beziehen sich auf Geschlechtsunterschiede. Im speziellen Vorurteile, welche sich auf Leistungsunterschiede zwischen den Geschlechtern beziehen, halten sich trotz Gleichberechtigung, Emanzipation und Förderprogrammen hartnäckig. Eines dieser Vorurteile bezieht sich auf die Leistungen von weiblichen Personen im Bereich der Mathematik. So sollen Männer im Allgemeinen bessere Leistungen in mathematischen Bereichen erzielen, als Frauen. Dabei legen einige Studien nahe, dass die Geschlechterunterschiede im mathematischen Bereichen gering sind (Hyde, Lindberg, Linn, Ellis und Williams, 2008) und dennoch sind Frauen in wissenschaftlichen, technischen und mathematischen Berufen weiterhin stark unterrepräsentiert.

Die Ergebnisse aus den PISA Studien 2012 zeigen wiederum deutliche Leistungsunterschiede zwischen 15 Jährigen Mädchen und Jungen unter anderem auch im Bereich der Mathematik. Ein Bericht aus dem Jahr 2013 der OECD (Organisation for Economic Co-operation and Development) spricht von „einem merklichen Leistungsunterschied zugunsten der Jungen“.

Eine der gängigsten Methoden des Primings ist die Präsentation einer bestimmten Information, wie beispielsweise „Dieser Test hat in der Vergangenheit Genderunterschiede gezeigt. Männer erreichten bessere Ergebnisse als Frauen“. Bereits eine solche Information kann dazu führen, dass bei Versuchspersonen ein Stereotyp aktiviert wird. Keller und Dauenheimer (2003) konnten zeigen das weibliche Teilnehmerinnen, welche die eben genannte Information erhielten, in einem Mathematiktest schlechter abschnitten, als männliche Personen. Personen welche die Information erhalten hatten, dass der folgende Test keine

Genderunterschiede zeigen würde und die Leistung von männlichen und weiblichen Personen gleich ausfallen würde. Keller (2002) fand ebenfalls, dass weibliche Teilnehmer welche eine solche Information erhielten schlechter abschnitten, als weibliche Teilnehmer, welche keinerlei Informationen erhalten hatten.

Eine andere weit verbreitete Methode des Primings ist die Vorstellung eines Artikels, welcher sich mit Genderthemen beschäftigt. In einer Studie von Dar-Nimrod und Heine (2006), wurden 220 Frauen fingierte Artikel mit vermeintlich wissenschaftlichen Erkenntnissen zwecks Lektüre ausgehändigt. In diesen Artikeln wurden die schlechteren mathematischen Leistungen von Frauen entweder mit genetischen, oder mit erfahrungsbedingten Gründen erklärt. Glaubte eine Frau in dem darauf folgenden Mathe-Quiz an die genetische Begründung, so schnitt sie im Test schlechter ab. Ging sie hingegen von der erfahrungsbedingten Erklärung aus, so erzielte sie ein besseres Ergebnis.

Neben den Stereotype Threat existiert ebenfalls die Bedingung des Stereotype Lift Effekts, welcher ebenfalls in Studien nachgewiesen werden konnte. Personen, die unter einem Stereotype Lift Effekt stehen, erbringen bessere Leistungen als Personen, welche nicht unter dem Effekt stehen (Walton & Cohen, 2003). Johnson, Barnard-Brak, Saxon, & Johnson (2012) stellten Stereotype Threat und Stereotype Lift Phänomene einander gegenüber. Sie verwendeten als Prime direkte Informationen mit denen den Versuchspersonen mitgeteilt wurde, dass erwartet wurde das Männer bzw. Frauen in dem bevorstehenden Test besser abschneiden würden. Frauen, welche der Stereotype Lift Bedingung ausgesetzt waren, schnitten besser in den Tests ab, als jene welche der Stereotype Threat Bedingung ausgesetzt waren.

In Unserer Studie beschäftigten wir uns sowohl mit dem Konzept des Stereotype Threats, als auch mit dem Phänomens des Stereotype Lift Effekts. Hierbei wollten wir eine neue Herangehensweise an diese Phänomene erarbeiten. Mit einer Baseline Studie sollte zunächst die Grundfähigkeit in Mathematik aller Versuchspersonen ermittelt werden, um sie sodann mit ihrer Fähigkeit unter dem Einfluss von einer Stereotype Lift oder Threat Bedingung zu vergleichen. Da wir

von jeder Versuchsperson die Grundfähigkeit in der für uns in diesem Zusammenhang relevanten Fähigkeit erhoben, hielten wir es für vertretbar auf eine Kontrollgruppe zu verzichten. Durch die zufällige Zuordnung der Personen zu den Versuchsgruppen, sollte die Bedingung der Randomisierung realisiert werden. Laut Literatur relevante Moderatorvariablen, wie Mathematiknote, persönliche Wichtigkeit von Mathematik, sowie Arbeitsstil und Motivation, wurden im Rahmen der Untersuchung ebenfalls erhoben.

1 Theoretischer Teil

1.1 Begriffsdefinition Priming

„Priming ist die (indirekte) Voraktivierung eines Gedächtnisinhalts durch einen Hinweisreiz“ (Herkner 2003, S. 167). Im „*Lexikon für Psychologie*“ wird Priming als „eine *assoziative Bahnung* bzw. automat. *Voraktivierung* eines mentalen Prozesses, v. a. einer bestimmten semant. Klasse im Wissenssystem“ (Städtler, 2003, S. 820) beschrieben. Ein von Herkner (2003) angeführtes Beispiel für Priming ist: Wenn Versuchspersonen ein Wort wie „grün“ dargeboten wird, so wird daraufhin nicht nur der Grün-Knoten aktiviert, sondern es werden gleichzeitig auch andere Assoziationen wie Wiese, oder andere Farben aktiviert. Ein einfaches Experiment, welches sich der Methode des Primings bedient, ist die „lexikalische Entscheidungsaufgabe“. Versuchspersonen sollen nach der Darbietung eines Reizes so schnell wie möglich entscheiden, ob es sich bei diesem Reiz um ein Wort handelt oder nicht. Wenn ein semantisch benachbartes Wort als Priming-Reiz dargeboten wird, gelingt diese Entscheidung schneller (Herkner, 2003). Zum Beispiel könnte der Priming-Reiz „rot“ dazu führen, dass das Wort „Tomate“ schneller als solches erkannt wird. Ist ein Priming-Reiz ein entferntes Wort, kann dies auch zu einer Urteilsverzögerung führen. Dies tritt vor allem ein, wenn die Versuchsperson einen Hinweisreiz erwartet, welcher mit dem Zielwort verwandt ist (Herkner 2003).

Fazio, Sanbonmatsu, Powell und Kardes (1986) zeigten in einem Experiment, dass Einstellungen welche schnell zugänglich sind, eher spontan aktiviert werden, als weniger zugängliche. In Fazio et al.'s (1986) Experiment wurden Personen mittels einer Vorstudie individuell mit Hilfe einer Reaktionszeitmethode untersucht. Für die individuellen Personen, welche an dieser Vorstudie und an den darauf folgenden Experimenten teilnahmen, wurden schnell zugängliche positive, sowie negative Einstellungsobjekte und schlecht zugängliche positive wie negative Einstellungsobjekte ermittelt. Ein Beispiel welches von Fazio, Sanbonmatsu, Powell und Kardes (1986) angeführt wurde ist, dass wenn eine Versuchsperson das Einstellungsobjekt „Vodka“ als positiv bewertet, sollte die Präsentation des

Wortes Vodka als Prime eine positive Evaluierung aktivieren bzw. anregen. In Zusammenhang mit den Untersuchungen von Fazio et al. (1986) hieße das dann, dass wenn ein präsentierte Zieladjektiv positiv ist, könnte es der Versuchsperson relativ schnell gelingen zu bemerken, dass das Zielwort eine positive Konnotation besitzt. Falls dies der Fall ist, sollte es der Versuchsperson leichter fallen eine Entscheidung zu treffen. Ähnlich sieht es auch mit negativen Worten als Prime aus. Das in diesem Fall angeführte Beispiel von Fazio et al. (1986) ist die Verbindung zwischen dem Prime Wort „Küchenschabe“ und dem Zieladjektiv „ekelig“. Als Baseline wurden Buchstabenketten (z.B. „BBB“) verwendet. Die Ergebnisse der Untersuchungen zeigten, dass sowohl starke positive, als auch starke negative Priming-Reize eine Reaktionsbeschleunigung zur Folge hatten, schwache Priming-Reize dagegen hatten keinen nachweisbaren Effekt.

Solche Primingprozeduren lassen sich auch auf viele weitere Begriffe ausweiten und können auch unser Verhalten beeinflussen. In einer Studie von Bargh, Chen und Burrows (1996) wurden Versuchspersonen mittels eines scrambled sentence tests geprimt. In den Sätzen waren einige Worte vertauscht, welche von den Teilnehmern an die richtigen Stellen gesetzt werden mussten. Eine Gruppe bekam Sätze, welche Begriffe enthielten, die sich mit dem Thema „altern“ beschäftigten. Dann wurde die Zeit ermittelt, welche die geprimte Versuchsgruppe sowie eine Kontrollgruppe brauchten, um einen Flur entlang zu gehen. Jene Gruppe, welche sich mit den Sätzen beschäftigten die „altern“ als Thema hatten, gingen signifikant langsamer den Flur entlang, als Personen in den Kontrollgruppen. Die Erklärung für dieses Phänomen sahen Bargh, Chen und Burrows (1996) in der automatischen Aktivierung des Stereotyps. Daher könnten auch bei der Nennung, oder dem Bewusstmachen des eigenen Geschlechts bestimmt Kognitionen und dazugehörige Assoziationen aktiviert werden.

Ob ein Priming wie erwartet wirkt, scheint auch von dem Umstand abzuhängen, dass die Teilnehmer sich der Beeinflussung durch das Priming selbst nicht bewusst werden. Lombardi, Higgins und Bargh (1987, zitiert nach Bargh et al., 1996, S. 237) zeigten, dass sie genau entgegengesetzte Effekte hinsichtlich der Beurteilungen der Versuchspersonen erhielten, wenn die Personen sich des Primings bewusst waren vs. jener Personen welche sich des Primings nicht

bewusst waren. Oder anders ausgedrückt, jene Versuchspersonen die sich des Primings bewusst waren reagierten genau entgegengesetzt zu dem gewünschten Effekt. In vielen Studien wird genau aus diesem Grund eine Kontrollfrage gestellt. Meistens wird diese nach dem Muster „Glauben Sie, dass [Priming] einen Einfluss auf [die weiteren Aufgabenstellungen in der Testung] hatte?“ gestellt.

Den Zeitraum in dem man Versuchspersonen einem Primereiz aussetzt, oder sie bittet sich mit diesem zu beschäftigen, kann variiert werden. So wurde bereits mit sehr geringen Primingzeiten Effekte erzielt werden. Murphy und Zajonc (1993) setzten ihre Versuchspersonen einem Prime (fröhliches oder wütendes Frauen/Männergesicht) für drei Millisekunden aus und baten sie chinesische Schriftzeichen auf einer Skala von 1 (gefällt mir gar nicht) bis 5 (gefällt mir sehr) zu bewerten. Es zeigte sich eine Verschiebung der Bewertung der Schriftzeichen. Schriftzeichen, vor deren Bewertung ein fröhliches Gesicht eingeblendet wurde, wurden deutlich positiver bewertet als Zeichen den ein wütendes Gesicht voranging. Dijksterhuis und van Knippenberg (1996), baten ihre Testpersonen Beispielsweise, sich fünf Minuten mit damit zu beschäftigen, sich einen Fußball Hooligan vorzustellen, verschiedene Verhaltensweisen sowie Attribute des Aussehen und Life Styles einer solchen Person aufzuschreiben. Dijksterhuis und van Knippenberg (1996, S. 272) beschreiben Stereotype als „knifes that cut both ways“ und meinen damit, dass Stereotypenaktivierung nicht nur den Zugang zu mit dem Stereotype konsistenten Eigenschaften (Traits) erleichtert, sondern auch den Zugang zu inkonsistenten Eigenschaften verhindert oder erschwert.

1.2 Stereotype: Begriffsdefinition, Ursprung und Funktionen

Um einen kleinen Überblick über den Begriff des Stereotyps, dessen Ursprung und dessen Funktionen zu liefern, wird im folgenden Kapitel auf diese Themen eingegangen. Außerdem werden speziell Geschlechtsstereotype und verwandte Begriffe erklärt und voneinander abgegrenzt.

1.2.1 Begriffsdefinition Stereotyp

Der Begriff *Stereotype* wurde erstmals von Lippman (1922, zitiert nach Dröge 1967, S. 134) in seiner Arbeit *Public Opinion* verwendet. In seinem Verständnis wird der Begriff Stereotyp als „eine erkenntnis-ökonomische Abwehreinrichtung gegen die notwendigen Aufwendungen einer umfassenden Detailerfahrung“ definiert. Wendt (2015, S. 4) schreibt in ihrem Buch das Stereotype „wertende Urteile, gebildet von Individuen oder Gruppen sind“. Weiter werden Stereotype als „konstruierte oder verzerrte Bilder die dazu dienen der Komplexität von Menschen, Dingen oder Ereignissen entgegenzuwirken“ (Wendt, 2015, S. 4) beschrieben. Durch die Verwendung von Stereotypen bilden wir zum Beispiel anhand der optischen Erscheinung einer Person erste Urteile bzw. Vorurteile darüber, wie sich Personen verhalten und welche Einstellungen sie haben. So haben wir beispielsweise eine bestimmte Vorstellung davon, welche Charaktereigenschaften und Fähigkeiten eine typische Krankenschwester besitzt. Wir sehen sie vielleicht als freundlich, hilfsbereit und empathisch, doch nicht jede Krankenschwester wird genau unserem Bild von ihr entsprechen. Weiße, Männer, sowie andere Mehrheiten werden als normal, oder auch typisch in den meisten Sektoren der Gesellschaft angesehen (Miller, Taylor & Buck, 1991; Walton & Cohan, 2003). Es ist daher weniger wahrscheinlich, dass diese Gruppen Objekte von negativen oder positiven Stereotypen werden (Miller et al., 1991). Aronson, Lustina, Good, Keough, Steele und Brown (1999) fanden in ihren Studien, bei denen sie Testpersonen baten Stereotype über verschiedene Gruppen zu beschreiben, heraus, dass nicht einer ihrer Teilnehmer Stereotype bezüglich der intellektuellen Fähigkeiten von Weißen äußerte.

Herkner (2003, S. 493) beschrieb Stereotype als „Übergeneralisationen“. Sie seien „die kognitive Komponente der Vorurteile –das subjektive Wissen bzw. die Meinung über die Außengruppe“ (Herkner, 2003, S. 493). Nach Herkner (2003) sei jedoch nur die Struktur der Stereotype kognitiv bedingt, der Inhalt hänge sowohl von kognitiven Prozessen, als auch von Motivationsprozessen ab.

1.2.2 Entstehung von Stereotypen

Wie entstehen Stereotype? Schneider (2004) unterscheidet diesbezüglich zwei Ansätze. Den Kulturellen Ansatz, dem zu Folge Stereotype über die Kultur an ihre Mitglieder weitergegeben werden, sowie den Sozial Kognitiven Ansatz. Der Sozial Kognitive Ansatz offeriert eine Anzahl an Mechanismen, bei welchen Einstellungen, Vorurteile, Werte und Überzeugungen die Entstehung von Stereotypen beeinflussen können. Schneider (2004) sieht diesen Ansatz als sehr oberflächlich, da er keine Erklärung dafür liefert wo die Überzeugungen, Einstellungen, Vorurteile und Werte erworben werden und wieso diese einen so großen Einfluss auf unsere Kognitionen haben. Jedoch betont er auch, dass der Kognitive Ansatz eine potentielle wichtige Quelle unserer Stereotype anspricht, nämlich unsere Erfahrungen mit anderen Menschen. Stereotype könnten also durch unsere Erfahrungen mit anderen Personen geformt werden, jedoch sei zu beachten, dass unsere Erfahrungen mit manchen Personen sehr beschränkt ausfallen und dadurch auch durchaus fehlerhaft sein können. Schneider (2004) sieht die Argumentation, dass unsere Kultur uns generell mit den Inhalten für unsere Stereotype versorgt und uns ermutigt manche Gruppen zu stereotypisieren und andere nicht, als zu simpel an. Er beschreibt uns als sowohl Erschaffer, als auch Produkt unserer Kultur. Die Kultur wird oft als Erklärungsansatz für die verschiedensten Dinge herangezogen ohne, dass wir eine hinreichende Erklärung haben, wie sie ihre Arbeit verrichtet. Wir verfügen laut Schneider (2004) bis jetzt nicht über präzise Modelle, wie kulturelle Lehren ihren Weg in unseren Verstand finden. Zu guter Letzt bleiben zwei große und diffuse Kategorien, welche dafür verantwortlich gemacht werden, warum manche Stereotype entstehen. Sie können Produkte unserer Kultur oder Subkultur oder unserer eigenen Erfahrungen sein (Schneider, 2004).

1.2.3 Funktionen von Stereotypen

Es können drei Ansätze zur Funktion von Stereotypen unterschieden werden. Ashmore und Del Bocha (1981, zitiert nach Snyder & Miene, 1994, S. 36) unterscheiden den Kognitiven Ansatz, den Psychodynamischen Ansatz und den Soziokulturellen Ansatz. Der Kognitive Ansatz geht davon aus, dass Menschen

nur eine limitierte Menge von Informationen verarbeiten können und Stereotype uns helfen, die kognitive Last zu reduzieren. Von der funktionalen Perspektive aus können Stereotype einen Betrag zur „cognitive economy“ leisten (Snyder & Miene, 1994, S.36). Außerdem schreiben Snyder und Miene (1994) in diesem Zusammenhang, dass Stereotype es ihren Besitzern ermöglichen, eingehende Informationen auf kontrollierbare Größe zu reduzieren und dadurch auch zu einem gewissen Grad die Vorhersehbarkeit der Welt zu vermitteln.

Der Psychodynamische Ansatz räumt den Stereotypen eine Reihe von Ego-Verteidigungsmechanismen ein. Beispielsweise sei der Mechanismus der Derogation anderer, insbesondere der Personen welche als Konkurrenten bezüglich limitierter Ressourcen gesehen werden, erwähnt (Taylor & Moriarty, 1987; zitiert nach Snyder & Miene, 1994, S. 36). Durch einen Vergleich nach unten mit Personen, welchen es noch schlechter geht als einem selbst, kann der Aufbau des eigenen Selbstwerts gefördert werden (Wills, 1981). Wenn beispielsweise ein männlicher Schüler in einem Mathematiktest eine relativ schlecht Note bekommt, könnte er seinen Selbstwert damit aufbauen, dass diese Note immer noch besser sei, als jene die der Großteil der Mädchen seiner Klasse bekommen hatte. Stereotype über andere Personen oder Personengruppen können also als nützliche Werkzeuge verwendet werden, um der nicht stereotypisierten Person oder Personengruppe ein besseres Selbstwertgefühl zu vermitteln (Snyder & Miene, 1994).

Der Soziokulturelle Ansatz sieht die Funktion von Stereotypen darin, dass sie Menschen helfen ihren Platz in ihren eigenen sozialen sowie kulturellen Gruppen zu finden und sich mit diesen zu identifizieren (Snyder & Miene, 1994). Durch die Sozialisation akquirieren Menschen Überzeugungen und Erwartungen über andere Gruppen und ihre Mitglieder. Da die Annahme dieser Stereotype oft den Umgang mit Personenaus der eigenen Gruppe erleichtert, stellen viele Leute diese auch nie in Frage (Snyder & Miene, 1994). Stereotype „functionally related to becoming a group member – to adopting the group values (norms) as the main anchorage in regulating experience and behavior“ (Sherif & Sherif, 1953, S.218; zitiert nach Snyder & Miene, 1994, S. 37). Welches bedeuten soll, dass Stereotype funktional damit verbunden sind, Teil einer Gruppe zu werden und das die

Adoption der Gruppenwerte wichtig für die Regulation der eigenen Erfahrungen und Verhaltensweisen ist. So erfüllen Stereotype auch eine soziale Funktion, indem sie uns helfen uns in unsere sozialen Gruppen einzufinden (Snyder & Miene, 1994).

1.3 Geschlechtsstereotype und Geschlechterrollen

„*Geschlechterstereotype* sind kognitive Strukturen, die sozial geteiltes Wissen über die charakteristischen Merkmale von Frauen und Männern enthalten“ (Ashmore & Del Boca 1979; Eckes 1997; zitiert nach Eckes, 2010, S. 178). Folgt man dieser Definition kommt man zu dem Schluss, dass Geschlechtsstereotype zum individuellen Wissensbesitz einerseits und andererseits den „Kern eines konsensuellen, kulturell geteilten Verständnisses von den je typischen Merkmalen der Geschlechter“ gehören (Eckes, 2010, S. 178). Eine weitere Definition stammt von Alfermann (1996, S. 31), wonach Geschlechtsstereotype „typische Eigenheiten von Männern und Frauen beschreiben und wie Wahrscheinlichkeitsannahmen wirken, die einen weiteren Handlungsverlauf steuern können.“

Wie Geschlechtsstereotype entstehen und warum so viele Interkulturelle Ähnlichkeiten bestehen ist eine Frage, die bereits viele Forscher beschäftigt hat. Alfermann (1996) schreibt diesbezüglich, dass beim Blick auf die Erwachsenenrolle bereits einige interkulturelle Gemeinsamkeiten gefunden werden. Als erstes nennt sie die unterschiedliche Macht, welche den Geschlechtern sowohl im öffentlichen, als auch im religiösen Leben zu Teil wird. So gehen in allen bekannten Nationen Männer als das dominante Geschlecht hervor und dies spiegelt sich auch in den Stereotypen Vorstellungen wieder. Eigenschaften wie „aggressiv“, „dominant“, „unabhängig“ und „unnachgiebig“ konnten in einer von Alfermann (1996, S. 20) angesprochenen Studie in allen 25 Nationen dem männlichen Stereotype zugeschrieben werden. Zweitens geht Alfermann (1996) davon aus, dass auch im privaten Leben ein solches Machtgefälle vorhanden ist. Als drittes nennt sie die Situation bei der Kinderbetreuung und Pflege, welche auch als eine typisch weibliche Aufgabe angesehen wird. Diese Rollenverteilung passt genau auf die gängigen Stereotype

und wird als solches auch an die nachfolgenden Generationen weitergereicht. Auch wenn viele Nationen sich mittlerweile um bessere Arbeitsverteilungen bemühen, wie beispielsweise sogenannte Frauenquoten oder Männerquoten in bestimmten Bereichen. Auch an mehr Möglichkeiten für Väter (wie beispielsweise das Elternjahr für Männer) wird gearbeitet, doch nicht alle nehmen diese Möglichkeiten auch wahr.

Die Rollen welche ihre Eltern ihnen vorleben, können von Kindern bereits früh eingeübt werden. Kinder können auch bereits sehr früh Geschlechter erkennen. Patterson und Werker (2002) wiesen bereits bei 8 Monate alten Kindern nach, dass diese mit einer hohen Treffsicherheit weibliche und männliche Stimmen und Gesichter richtig zuordnen konnten. Konzepte wie Maskulinität und Femininität entwickeln sich bei Kindern ebenfalls relativ früh. Martin, Wood und Little (1990) fanden in Studien Hinweise, dass Kinder zuerst Assoziationen bezüglich der Charakteristika ihres eigenen Geschlechts lernen und erst später jene des anderen Geschlechts. Außerdem wurden die stereotypen Bewertungen stärker, je älter die untersuchten Kinder waren. Auch die Präferenz für stereotypes Spielzeug konnte bei einer Studie von Robinson und Morris (1986) bereits bei zwei bis fünf Jährigen Kindern nachgewiesen werden. In jeder Alterskategorie zeigten männliche untersuchte Kinder ca. zu 75 % eine Präferenz für „Jungs Spielsachen“ (typische „männliche“ Spielsachen). Die untersuchten Mädchen zeigten erst ab dem Alter von fünf Jahren ähnliche Präferenzen.

Ein eng verwandter Begriff ist der der *Geschlechterrolle*. Geschlechterrollen werden teils präskriptiv „im Unterschied zu deskriptiv verstandenen Geschlechterstereotypen konzipiert (Alfermann, 1996), teils werden sie, insbesondere im Kontext der Analyse von Familien-, Berufs- und Führungsrollen, sowohl in ihren deskriptiven, als auch in ihren präskriptiven Funktionen diskutiert“ (Eagly & Karau 2002; Eagly, Wood & Diekman 2000; zitiert nach Eckes, 2010, S. 178). Eckes (2010, S. 178) schrieb, dass die Betonung bei diesem Konzept auf den „sozial geteilten Verhaltenserwartungen“ liegt, welche sich an Individuen aufgrund ihres sozial zugeschrieben Geschlechts richten. Geschlechtsrollen entwickeln sich durch den Einfluss von drei zentralen Faktoren: Biologische Faktoren, Umgebungsfaktoren und Kognitive Faktoren. Diese Faktoren

beeinflussen sich gegenseitig, jedoch gibt es Ansätze welche sich auf den besonderen Einfluss eines Faktors konzentrieren (Lewis & Weinraub, 1979). In jedem Fall entwickeln Menschen eine Vorstellung davon, was das für ihr Geschlecht angemessenes Verhalten ist.

Bezüglich der Geschlechtstypisierung¹ gibt es verschiedene Ansätze, von denen im Folgenden einige kurz angeführt werden. Der psychoanalytische Ansatz geht auf Sigmund Freud (1856 – 1939) zurück. Er nahm an, dass die erste Geschlechtstypisierung im Alter von ungefähr drei bis sechs Jahren stattfindet, in der sogenannten phallischen Phase. Die Geschlechterrollenübernahme nimmt vor allem durch die sexuelle Orientierung Gestalt an. Durch sexuelle Gefühle dem gegengeschlechtlichen Elternteil gegenüber gerät das Kind in eine Rivalität mit dem gleichgeschlechtlichen Elternteil, welches mit dem Ödipuskomplex einhergeht. Allerdings wird dieser Ansatz von vielen als überholt angesehen, da von Freud die Geschlechtsrollenentwicklung nur von dem Verhalten der Eltern abhängig gemacht wurde. Außerdem konzentrierte sich sein Ansatz in erster Linie darauf, den Ursprung für psychische Probleme im Erwachsenenalter zu finden (Berk, 2011).

Ein weiterer Ansatz ist der, der Sozialen Lerntheorie, welcher vor allem davon ausgeht, dass Kinder Verhalten von Modelpersonen beobachten und imitieren. Durch Albert Bandura (1971, zitiert nach Berk, 2011, S. 20) wurde dies als Modellernen oder Beobachtungslernen bekannt. Die Geschlechtsidentität eines Kindes entsteht nach der sozialen Lerntheorie durch Selbstwahrnehmung. Kinder organisieren erst später geschlechtstypische Verhaltensweisen, welche es durch

¹ „Geschlechtstypisierung meint in der Psychologie den Erwerb einer traditionell männlichen oder weiblichen Rolle, die in der Regel durch Lernen stattfindet.“ (Stangl, 2011, WWW: <http://lexikon.stangl.eu/3568/geschlechtstypisierung>)

Beobachtung und Nachahmung gelernt hat, zu geschlechtsbezogenen Vorstellungen über sich selbst (Berk, 2011).

Trautner (1991) unterscheidet zwei Ansätze der sozialen Lerntheorie: die Bekräftigungstheorie und die Imitationstheorie. Die Bekräftigungstheorie nimmt an, dass geschlechtsangemessenes Verhalten bereits von Geburt an verstärkt wird und geschlechtsunangemessenes Verhalten entweder ignoriert, oder sogar bestraft wird. So werden Mädchen beispielsweise dafür gelobt, wenn sie im Haushalt mithelfen wollen. Dagegen könnten Burschen seltsame Blicke von Eltern und Mitschülern erhalten, wenn sie eine Vorliebe für Puppen entwickeln (Trautner, 1991). Die Imitationstheorie vertritt den Standpunkt, dass Kinder vor allem durch das Beobachten und Nachahmen von gleichgeschlechtlichen Modellen geschlechtsspezifisches Verhalten erlernen. Alfermann (1996) argumentiert, dass die soziale Lerntheorie dem Verhalten selber eine zu große Rolle beimisst und dabei andere für die Geschlechtsrollenentwicklung wichtige Aspekte, wie etwa Kognitionen und äußere Merkmale außer Acht lässt. Die kognitive Entwicklungspsychologie geht davon aus, dass (anders als bei der sozialen Lerntheorie) die Selbstwahrnehmung vor dem Verhalten stattfindet. So kommt das Kind in den Kindergartenjahren zu der Erkenntnis, dass sein Geschlecht unveränderbar ist, es entwickelt Geschlechtskonsistenz. Geschlechtskonsistenz kann verstanden werden als „die gefestigte Erkenntnis, dass das Geschlecht biologisch festgelegt und permanent ist, also auch dann dasselbe bleibt, wenn Kleidung, Frisur oder spielerische Aktivitäten sich ändern“ (Berk, 2011, S. 371). Kohlberg (1966, zitiert nach Trautner 1991, S. 384) geht davon aus, dass die Geschlechtstypisierung sich in drei Schritten abläuft. In einem ersten Schritt erkennt das Kind seine Geschlechtszugehörigkeit, welches mit zwei bis drei Jahren geschieht. Das Kind orientiert sich bei der Kategorisierung am Anfang an Merkmalen äußerer Ähnlichkeiten zwischen sich selbst und anderen Personen. Bei dem zweiten Schritt entwickeln sich geschlechtsbezogene Wertungssysteme und Einstellungen. Diese gehen im Regelfall mit der Bevorzugung des eigenen Geschlechts einher. „Kohlberg nimmt an, dass Kinder, sobald sie von der Unveränderlichkeit ihrer Geschlechtszugehörigkeit überzeugt sind, sich hinsichtlich ihres Verhaltens und seiner Bewertungen vorrangig am Kriterium der

Geschlechtsangemessenheit orientieren“ (Trautner, 1991, S. 384). Die Entwicklungsprozesse die zu dem dritten Entwicklungsschritt, nämlich der Bindung und selektiven Nachahmung, wie auch Identifikation mit gleichgeschlechtlichen Personen führen, sind folgende: „1. Es werden weitere Informationen über geschlechtsbezogener Merkmale gesucht; 2. Was die wahrgenommene Geschlechtsdifferenzierung bestätigt wird positiv bewertet; 3. die Übereinstimmung mit geschlechtsangemessenen Standards gewinnt verhaltenssteuernde Bedeutung“ (Trautner, 1991, S. 384). Bei der kognitiven Theorie kritisiert Alfermann (1996), dass ausschließlich kognitive Prozesse bei der Geschlechtsrollenentwicklung Beachtung finden und soziale Einflüsse nicht berücksichtigt werden.

1.4 Das Dissoziative Modell nach Devine

Das Dissoziative Modell nach Devine (1989) ist ein kognitives Modell, welches sich mit dem Umgang von Stereotypeninformationen beschäftigt. Devine geht von einem Zwei Stufen Modell der kognitiven Verarbeitung aus. Bei diesem Modell wird davon ausgegangen, dass durch automatische Verarbeitung Informationen aktiviert werden (in diesem Fall die Stereotype) und durch die bewusste (kontrollierende) Verarbeitung diese Information sodann angefochten, akzeptiert oder ignoriert werden können. Devine (1989) untersuchte in ihren drei Studien was passiert, wenn diese automatische Informationsverarbeitung in Kraft tritt.

In mehreren separaten Studien untersuchte Devine (1989) hellhäutige Psychologiestudenten, unter anderem mit Hilfe einer Rassismus Skala und einer besonderen Primingmethode. In der ersten Studie dieser Reihe bat sie weiße Psychologiestudenten, Komponenten des gängigen Stereotypes über dunkelhäutigen Personen auf ein liniertes Blatt Papier zu schreiben. Die Teilnehmer wurden explizit darauf hingewiesen, dass sie nicht nach ihrer eigenen Meinung und oder Einstellungen zu dunkelhäutigen Personen gefragt wurden, sondern nach Komponenten des allgemeinen kulturellen Stereotyps über Schwarze. Im Anschluss wurde den Studenten der Rassismusfragebogen bzw. die Rassismusskala vorgegeben. Die Ergebnisse unterstützen die Annahme ihres Modells, dass sowohl Personen mit einem hohen, als auch niedrigem

Vorurteilslevel, das gleiche Wissen über die üblichen kulturellen Stereotype über farbige Personen besitzen. In einer zweiten Studie von Devine (1989) wurde eine große Gruppe von Studentinnen mittels der Rassismusskala, je nach ihrem erreichten Punktwert, in Personen mit hohen Werten in dem Test (und somit einem hohen Vorurteilsniveau) und Personen mit niedrigen Punktwerten (Personen mit niedrigem Vorurteilsniveau) eingeteilt (Personen im mittleren Bereich wurden ausgeschlossen). Nach dem Durchlaufen weiterer Auswahlkriterien, wurden die verblieben Testpersonen einem Priming unterzogen. Die Priming Aufgabe, der automatischen und bewussten Verarbeitung nach Bargh und Pietromonaco (1982), bestand aus der Präsentation von mit farbigen Menschen assoziierten Worten, wie Beispielsweise „faul, athletisch, arm“ welche kurzzeitig (80 ms) präsentiert wurden, danach folgte die Präsentation einer Maske (ungeordnete Buchstaben), außerdem wurde ein gewisser Anteil neutraler Wörter eingefügt. Dadurch, dass diese Worte so kurz gezeigt wurden, ging Devine (1989) davon aus, dass automatische Verarbeitungsprozesse in Kraft treten würden. Nach diesem Priming lasen die Teilnehmer eine Geschichte über „Donald“, dessen ethnische Zugehörigkeit nicht erwähnt wurde. In dieser Geschichte wurde Donald nicht näher beschrieben und er tat einige Dinge, welche man positiv oder negativ interpretieren konnte. Nach dem lesen dieser Geschichte wurden die Teilnehmer gebeten, ihre Eindrücke über diese Person „Donald“ wiederzugeben und ihn auf einer Skala von 0 bis 10 anhand verschiedener Traits² einzuschätzen. Sowohl Personen mit einem niedrigen, als auch hohem Vorurteilsniveau schätzten „Donald“ als sehr Feindselig auf den dazugehörigen Skalen ein. Devine (1989)

² „Ein Trait bezeichnet in der Psychologie einen Persönlichkeitszug, also ein für einen bestimmten Menschen typisches Verhaltens-oder Veranlagungsmuster, das sich in seiner Art zu fühlen und zu handeln ausdrückt. Diese Muster können durch Fragebögen zur Erhebung der Selbst-und der Fremdeinschätzung erfasst werden.“ (Stangl, 2011, WWW: <http://lexikon.stangl.eu/4752/trait/>).

sieht hierin ein Beweis dafür, dass dieses unbewusste Priming des ethnischen Stereotyps eine Verbindung zwischen Schwarzen und Feindseligkeit aktiviert. In einer weiteren Studie konnte Devine (1989) nachweisen, dass nur Personen mit einem geringen Vorurteilsniveau die automatisch aktivierten stereotypkonformen Gedanken hemmten und diese durch reflektierte Überlegungen der Gleichheit und Negation des Stereotyps ersetzen. Diese Befunde legen nahe, dass es Personen mit einem hohen Vorurteilsniveau schwerer fällt, außerhalb ihrer Stereotype zu denken und über diese zu reflektieren.

Mitulla (1997) fasst in ihrem Buch *Die Barriere im Kopf: Stereotype und Vorurteile bei Kindern gegenüber Ausländern* die sechs zentralen Aussagen des Dissoziations Modells nach Devine (1989) wie folgt zusammen:

- „1. Stereotype sind gelernte Sets von Assoziationen, die durch gelernte Hinweisreize automatisch aktiviert werden können.
2. Vorurteilsvolle und vorurteilslose Menschen unterscheiden sich nicht im Grad der Kenntnis der kulturellen Vorurteile und Stereotype einer Gruppe, da sie diese als Mitglieder einer Gesellschaft im Laufe ihrer Sozialisation erlernt haben.
3. Vorurteilsvolle und Vorurteilslose haben unterschiedliche Überzeugungen. Vorurteilslose Personen lehnen die im kulturellen Stereotyp enthaltenen Zuschreibungen weitgehend ab. Bei diesen Personen hat ein Reflektionsprozess stattgefunden, der zu einer bewussten Ablehnung der stereotypen Zuschreibung führte. Bei vorurteilsvollen Personen entsprechen die ethnischen Vorurteile weitgehend ihren Überzeugungen.
4. Die einmal erworbenen kulturellen Stereotype und Vorurteile bleiben weiterhin im Gedächtnis präsent, auch wenn sie nicht mehr der aktuellen Einstellung entsprechen.
5. In Situationen, in denen vor allem automatische Informationsverarbeitungsprozesse ablaufen, können so Stereotype unabhängig vom Vorurteilsniveau die Wahrnehmung und Beurteilung von Situationen und Personen beeinflussen.

6. In Situationen, in denen vor allem bewusste Informationsverarbeitungsprozesse stattfinden, reagieren Personen entsprechend ihren persönlichen Überzeugungen.“ (Mitulla, 1997, S. 66 - 67)

Ähnlich wie in Devines (1989) Modell, anhand von Stereotypen über Afroamerikaner beschrieben, könnten Personen auch auf andere gängige Stereotype reagieren. Wir würden annehmen, dass das für uns in unserer Untersuchung relevante Stereotyp ebenfalls in unserer Zielpopulation verbreitet ist. In den Berichten wurde das gängige Stereotyp als wahrheitsgemäß dargestellt, jedoch ohne auf Besonderheiten einzugehen. In wie fern die einzelnen Teilnehmer selbst mit diesem Stereotyp übereinstimmen ist uns bei unserer Untersuchung im Vorhinein natürlich nicht bekannt, jedoch wird auch dieser Frage nachgegangen.

1.5 Geschlechtsunterschiede bei mathematischen Fähigkeiten

Geschlechtsunterschiede in mathematischen und technischen Bereichen sind ein gängiges Stereotyp, welches sich hartnäckig hält. Hyde, Fennema und Lamon (1990) untersuchten den Wahrheitsgehalt dieses dominanten Stereotyps in einer groß angelegten Meta-Analyse. In den Alterskategorien von fünf bis zehn Jahren und 11 bis 14 Jahren zeigte sich ein geringer Geschlechtsunterschied, der darauf hinauslief das Mädchen bessere Ergebnisse erreichten. Wieso Mädchen im Alter von 11 bis 14 bessere Ergebnisse erreichen als männliche Personen, wird von Steele (1997) mit dem Prozess des Stereotype Stratification erklärt. Bei diesem Prozess sehen sich Teilnehmer, welche von dem negativen Stereotyp betroffen sind, selbst als Teil einer Subgruppe auf die dieser negative Stereotyp nicht zutrifft. Die weiblichen Teilnehmer schätzten, dass Männer Mathematik mehr mögen und besser darin seien als Frauen, allerdings schätzten sie auch, das Mädchen und Jungs sich in diesen Bereichen nicht unterscheiden würden. Außerdem zeichneten die Mädchen eher einen Mann als eine Frau, wenn ihnen eine Geschichte über einen erwachsenen Mathematiker erzählt wurde, jedoch zeichneten sie häufiger ein Mädchen, wenn die Geschichte von einem kindlichen Mathematiker handelte.

In der Alterskategorie von 15 bis 18 und darüber, verschob sich der zuvor beschriebene Unterschied deutlich zum Vorteil der männlichen Versuchspersonen

(Hyde, Fennema & Lamon, 1990). Dies war einer der Gründe uns für eine Untersuchungsgruppe von Kindern im Alter von 14 bis 17 Jahren zu entscheiden. Hyde et al. (2008) fand in einer weiteren Studie mit Daten aus insgesamt 10 repräsentativen Staaten deutliche Hinweise dafür, dass keinerlei Geschlechtsunterschiede messbar waren. Wir erwarten daher ein ähnliches Bild bei unserer Studie.

Eine Quelle für die Unterschiede in den mathematischen Leistungen im Vergleich zwischen verschiedenen Ländern, Geschlechtern und ähnlichem bietet die PISA Studie. In dem Bericht der Organisation for Economic Co-Operation and Development (OECD) aus dem Jahr 2013 wird behauptet, die PISA Studie 2012 würde das bislang umfassendste Bild der in den Schulen vermittelten Mathematikkompetenzen liefern. Es wird untersucht, in wie fern sich Schüler in verschiedenen Bereichen der Mathematik Wissen aneignen und wie gut sie dieses Wissen einzusetzen vermögen. Die Ergebnisse der PISA Studie von 2012 zeigen, laut einer Zusammenfassung des OECD (2013) Berichts, deutliche Unterschiede der Leistungen der 15 Jährigen im Bereich Mathematik im Vergleich der einzelnen Länder. Außerdem steht im Bericht zum Thema Geschlechtsunterschiede:

„Die PISA-Ergebnisse zeigen, dass **Mädchen selbst dann, wenn sie ebenso gute Mathematikleistungen erbringen wie Jungen**, laut eigenen Angaben im Durchschnitt über weniger Ausdauer, eine weniger positive Einstellung zum Lösen von Problemen, eine geringere intrinsische und instrumentelle Motivation für das Lernen in Mathematik, sowie ein geringeres Vertrauen in die eigenen Lernfähigkeiten in Mathematik verfügen und größere Angst vor Mathematik haben, als Jungen; sie führen ein Versagen in Mathematik darüber hinaus mit größerer Wahrscheinlichkeit auf sich selbst zurück, anstatt auf äußere Faktoren.“ (OECD Bericht, 2013, S. 18)

Der Bericht geht weiter darauf ein, dass Mädchen in 37 der 65 teilnehmenden Länder schwächere Leistungen zeigten als Jungen. Außerdem lagen in den meisten Ländern die Ergebnisse der leistungstärksten Mädchen in Mathematik unter denen der leistungstärksten Jungen. Die geschlechtsspezifischen

Unterschiede in Bezug auf die Einsatzbereitschaft, die Motivation und das Vertrauen in die eigenen Fähigkeiten, werden als „besorgniserregend“ beschrieben. Zu bemerken sei jedoch, dass die Unterschiede innerhalb der Geschlechtergruppen noch deutlich größer waren, als zwischen den Geschlechtern und dass die geschlechtsspezifischen Leistungsunterschiede erheblich zwischen den einzelnen Ländern variierten. Dies sehen die Autoren des OECD Berichts als Hinweis darauf, dass „Stärken und Schwächen in bestimmten Schulfächern nicht naturgegeben sind, sondern erworben und oftmals gesellschaftlich verstärkt werden“ (OECD, 2013, S. 23). Eine nicht ganz unbegründete Annahme. In einer Studie zur „Rolle der Genderstereotype von Müttern“ wurden von Tomasetto, Alparone und Cadinu (2011) Hinweise darauf gefunden, dass die Einstellung (Befürwortung oder Ablehnung) bezüglich solcher Stereotypen einen Einfluss auf die Leistungen der untersuchten Mädchen hatten, wenn diese mit den Stereotypen konfrontiert wurden.

Um den gängigen Stereotypen entgegenzuwirken wurde in der Zusammenfassung des PISA Berichts 2012 vorgeschlagen Mädchen Mathematik „auf interessante Weise näher zu bringen, geschlechtsspezifische Stereotypen in den Schulbüchern zu identifizieren und zu beseitigen, weibliche Rollenmodelle zu fördern und Lernmaterialien zu verwenden, die Mädchen ansprechen“ (OECD, 2013, S. 23). Außerdem sollte man Eltern, Lehrer sowie die Gesellschaft, als Ganzes in die Bemühungen miteinbinden, gegen solche Stereotypenvorstellungen über die Leistungsmöglichkeiten von Mädchen vorzugehen.

1.6 Stereotype Threat

Der Begriff des Stereotype Threat (ST) wurde von Steele und Anderson (1995, S. 797) als „being at risk of confirming, as self-characteristic, a negative stereotype about one's group.“ beschrieben. In ihren Studien konzentrierten sich Steele und Anderson (1995) darauf, das in den USA weit verbreitete negative Stereotyp bezüglich der intellektuellen Fähigkeiten von Afroamerikanern zu untersuchen. Der verwendete Test bestand aus als schwierig angesehenen Teilen des Graduate Record Examination (GRE). Der GRE ist ein standardisierter Test, den StudentInnen in den USA zur Aufnahme eines Doktoratsstudiums absolvieren

müssen. Als Prime verwendeten sie in den ersten drei Experimenten die Beschreibung des von den Versuchspersonen auszufüllenden Tests. Der Test wurde als ein diagnostisches Werkzeug zur Ermittlung der intellektuellen Fähigkeiten beschrieben. In einem weiteren Experiment dieser Reihe (Experiment vier) wurden die Teilnehmer gebeten ihre Ethnie anzugeben. Die Autoren gingen davon aus, dass durch diese Maßnahmen die Stereotypen bezüglich der intellektuellen Fähigkeiten aktiviert werden würden. Den jeweiligen Kontrollgruppen wurde mitgeteilt, dass der Test lediglich die Problemlösungsfähigkeit (Experiment eins, zwei und drei) der Probanden messen würde, bzw. sie mussten ihre Ethnie nicht angeben (Experiment vier). Dunkelhäutige Personen, welche der Stereotype Threat Bedingung zugeteilt wurden, schnitten schlechter ab, als weiße und dunkelhäutige Personen in den Kontrollgruppen. Zwischen den Ethnien in den Kontrollgruppen zeigten sich keine Unterschiede. Die vorgefunden Ergebnisse bezüglich der reduzierten Leistungen der geprimten afroamerikanischen Teilnehmer erklärten sich die Autoren durch eine Reduktion der Geschwindigkeit und der Genauigkeit während der Testbearbeitung. Steele und Anderson (1995) argumentieren, dass die Existenz eines negativen Stereotyps bezüglich einer Gruppe der man angehört bedeutet, dass in Situationen in denen dieser Stereotyp zutreffend ist, das Risiko besteht diesem Stereotyp als Selbstcharakterisierung, einerseits für sich selbst und andererseits für andere, denen dieser Stereotyp bekannt ist, zu entsprechen. Wenn der Stereotyp etwas so selbstrelevantes wie intellektuelle Fähigkeiten enthält, kann diese Bedrohung so gravierend sein, dass sie die intellektuellen Fähigkeiten beeinträchtigt. Steele und Anderson (1995, S. 797) schrieben außerdem: "When the allegations of the stereotype are importantly negative, this predicament may be self-threatening enough to have disruptive effects of its own." Durch eine Reihe an Mechanismen löst Stereotype Threat daher *self-threat* aus, dies kann wiederum dazu führen, dass die intellektuelle Funktionsfähigkeit der betroffenen Personen beeinträchtigt wird. Dieser Effekt kann vor allem bei standardisierten Tests beobachtet werden (Steele & Aronson, 1995).

In dieser Tradition gibt es eine Vielzahl an Untersuchungen, welche an diese Ergebnisse anknüpfen. So beschränkt sich das Phänomen des Stereotype Threat

nicht nur auf Ethnien, welche negativ stereotypisiert werden, auch andere Minoritäten weisen unter bestimmten Umständen eine Leistungsverminderung unter dem Einfluss eines Stereotype Threat auf. Es wurden eine große Menge an gängigen Stereotypen auf unterschiedlichste Weisen geprimt und die Wirkung auf die betroffenen Personen untersucht. So wurde beispielsweise von Abrams, Crisp Marques, Fagg, Bedford und Provias (2008) der negative Effekt von Stereotypen bezüglich älterer Leute auf deren Mathematikleistung untersucht. Sie konnten zeigen, dass generationsübergreifender Kontakt, oder die bloße Vorstellung dessen, in niedrigeren testbezogenen Ängstlichkeitswerten und besseren Testergebnissen resultierte.

1.6.1 Stereotype Threat im Zusammenhang mit Priming

Im Zusammenhang mit Priming wurden in Kapitel 1.5 viele verschiedene Studien erwähnt, in denen Priming verwendet wurde um ST (Stereotype Threat) und SL(Stereotype Lift) Effekte zu erzeugen. Marx und Stapel (2006) gehen in ihren Experimenten auf den Unterschied zwischen ST und Primingeffekten näher ein. In ihren Experimenten verwendeten sie einen Scrambled Sentence Test, welcher als Prime die Eigenschaft „dumb“ (bzw. dumm) aktivieren sollte, einen Mathematiktest (Teil des Graduate Record Exam) sowie einige Fragen, die sich mit dem Selbstkonzept der Teilnehmer bezüglich der Bestätigung des negativen Stereotyps beschäftigten. Der Stereotype Threat wurde hier mittels der Mittelung erzeugt, dass es sich um einen besonders schweren Mathematiktest handeln würde, das Trait Priming wurde mittels des Scrambled Sentence Tests erreicht. In ihren Experimenten konnten die Autoren zeigen, dass ST sich negativ auf die Testleistungen von Zielpersonen und deren Selbstkonzept (bzw. „Threat based concerns“) auswirkt, weil nur bei den Zielpersonen eine Verbindung zwischen dem Selbstkonzept und dem negativen Gruppenstereotyp bestand. Das Trait Priming wirkte sich sowohl auf Zielpersonen, als auch auf nicht Zielpersonen negativ aus, hatte jedoch keinen Effekt auf die Threat based concerns. Die Autoren argumentieren, dies sei ein Beweis dafür, dass das Trait Priming nicht auf eine Verbindung zwischen dem Gruppenstereotyp und dem sozialen Selbst angewiesen ist.

Ambady, Shih, Kim und Pittinsky (2001) konnten in ihrer Studie zeigen, wie sich positive und negative Stereotype auf die mathematischen Leistungen von verschiedenen Altersgruppen auswirken. Es konnte gezeigt werden, dass Stereotypmanipulation bereits bei Kindern im Alter von fünf bis sieben Jahren nachgewiesen werden konnte. Mädchen in den Altersstufen von fünf bis sieben und elf bis dreizehn Jahren, welche auf ihre asiatische Herkunft aufmerksam gemacht worden waren, schnitten in einem Mathematiktest besser ab als jene, welche auf ihre weibliche Identität aufmerksam gemacht worden waren. Außerdem zeigten sie, dass Mädchen in genanntem Alter gute mathematische Fähigkeiten eher einer männlichen, als einer weiblichen Person zu zuschreiben. Bei den weiblichen Teilnehmern im Alter von acht bis zehn zeigte sich ein anderes Bild. In dieser Altersgruppe schnitten Mädchen, welche auf ihr Geschlecht aufmerksam gemacht wurden am besten ab und schrieben gute mathematische Fähigkeiten eher weiblichen Personen zu, als einer männlichen. Ambady et al. (2001) erklären diese Ergebnisse damit, dass Kinder in diesem Alter sehr chauvinistisch sind und daher ihr eigenes Geschlecht als besser einschätzen. Sie berufen sich hierbei auf eine Studie von Powlishta (1995), in der Kinder im Alter von acht bis zehn, die die Eigenschaften ihres eigenen Geschlechts positiver einschätzten, als jene des anderen Geschlechts. Bei den männlichen Teilnehmern zeigte sich, dass männliche Probanden in den Alterskategorien fünf bis sieben sowie elf bis dreizehn die besten Ergebnisse erzielten, wenn sie auf ihr Geschlecht oder auf ihre asiatische Herkunft aufmerksam gemacht wurden. Auch sie schrieben gute mathematische Fähigkeiten eher einer männlichen Person zu. In der Alterskategorie von sechs bis zehn zeigten sich bei den männlichen Probanden die besten Ergebnisse, wenn ihre männliche Identität aktiviert wurde. Diese Befunde halfen uns dabei, uns für eine Altersgruppe für unsere Testung zu entscheiden.

In der Literatur finden sich verschiedene Priming-Methoden um Stereotype Threat und Stereotype Lift Effekte hervorzurufen. Bei jüngeren Kinder wurden beispielsweise von den Mädchen verlangt, eine Zeichnung eines weiblichen Kind, welches eine Puppe hielt, mit Farben auszumalen, während die männlichen Teilnehmer das Bild eines Jungen ausmalen sollten (Ambady et al., 2001). Andere

Methoden umfassen das Zeigen von stereotypen Werbungen (Ihme & Mauch, 2007), oder auch das Vorlesen von Geschichten und das anschließende Zeichnen der Hauptpersonen der Geschichten (Tomasetto, Alparone & Cadinu, 2011). Der Nimrod und Heine (2006) verwendeten Artikel mit verschiedenen Themen als Priming-Reize und konnten deutliche ST Effekte nachweisen. Frauen, welche einen Artikel mit dem Thema „Geschlechtsunterschiede in Mathematik sind genetisch bedingt“ bekamen und Frauen, welche einen Artikel bekamen, welcher sich nur mit Gender beschäftigte (ohne explizit auf mathematische Leistungen einzugehen), schnitten schlechter ab als jene, welche Artikel bekamen die besagten, dass Leistungsunterschiede in Mathematik auf Erfahrungen zurückzuführen seien, oder dass keine Geschlechtsunterschiede im mathematischen Bereich existieren. Appel, Kronberger und Aronson (2011) übersetzten diese Artikel für ihre Studien, die sie sich mit dem Effekt von ST auf die Aktivitäten, welche notwendig sind um Wissen überhaupt aufzubauen, beschäftigten. Da diese Methode sich als praktikabel und wirkungsvoll erwies, sowie als sehr geeignet für die Onlinestudie erscheint, wäre dies eine geeignete Priming Methode für unsere Studie.

1.6.2 Moderatoren für Stereotype Threat

In Abschnitt 1.5.1 wurden bereits einige Studien im Zusammenhang mit Stereotype Threat besprochen. Für das Auftreten von Stereotype Threat sollten jedoch auch ganz bestimmte Rahmenbedingungen vorliegen. Zu den Moderatoren für Stereotype Threat gehören die Zugehörigkeit zu einer negativ stereotypisierten Gruppe, die Identifikation mit der negativ stereotypisierten Domäne und ein Test mit hohem Schwierigkeitsniveau (Martiny & Götz, 2011).

Um ein negativen Stereotyp zu aktivieren gibt es eine Vielzahl von Möglichkeiten, von denen bereits einige besprochen wurden. Den an der Untersuchung teilnehmenden Personen muss in jedem Fall, die eigene Unterlegenheit gegenüber einer außenstehenden Gruppe (daher einer Gruppe welcher die Versuchsperson nicht angehört) vermittelt werden, damit der typische leistungsreduzierende Effekt von Stereotype Threat zustande kommt (Aronson, Lustina, Good, Keough, Steele & Brown, 1999). Dabei kann die Beschreibung des

vorgegebenen Tests einen großen Einfluss auf, das Auftreten von Stereotype Threat Effekten haben. Huguet und Régner (2007) zeigten in einer Studie, dass die Beschreibung des Tests einen großen Effekt auf die Leistungen haben kann. Sie zeigten dies, indem sie den Probanden eine Aufgabe einmal als „Geometrie Test“ und einmal als *Gedächtnistest* vorstellten. Männliche Teilnehmer schnitten in der Geometrie und Weibliche in der Gedächtnistestbedingung besser ab.

Es konnte nachgewiesen werden, dass es nicht von entscheidender Bedeutung ist, ob die stereotypisierten Versuchspersonen glauben, dass die aktivierten und anwendbaren Stereotypen richtig sind, damit es zu einem negativen Einfluss auf die Leistung kommt (Steele, 1997; Steele & Aronson, 1995). Huguet und Régner (2009) konnten in ihrer Studie sogar zeigen, dass selbst wenn teilnehmende Personen nicht an das aktivierte Stereotype glaubten, dennoch eine Leistungsreduktion eintrat. Schmader, Johns und Barquissau (2004) fanden jedoch bei Frauen, welche ihr Hauptfach in mathematischen Bereichen hatten, aber dennoch Statusunterschiede zwischen den Geschlechtern als legitim ansahen und die gängigen Stereotype bezüglich der mathematischen Fähigkeiten von Frauen befürworteten, weniger Interesse daran ihre Studien in diesem Feld weiter zu führen. Außerdem waren Personen, welche den Stereotyp für legitim hielten auch anfälliger für den leistungsreduzierenden Effekt von Stereotype Threat, als jene Personen, welche dies nicht taten.

Die Zugehörigkeit zu einer negativ stereotypisierten Gruppe ist ebenfalls wichtig, damit eine Aktivierung des Stereotyps stattfinden kann und anwendbar ist. Natürlich gehört jeder nicht nur einer, sondern gleich mehreren sozialen Gruppen an und über die meisten dieser sozialen Gruppen kursieren verschiedene Stereotypen. Stereotype Threat kann daher auch jeden betreffen, wenn die „richtigen“ Stereotypen aktiviert werden (Steele, 1997; Aronson, 2004). Außer der Zugehörigkeit zu der stereotypisierten Gruppe, kann auch noch die Identifikation mit eben jener Gruppe wichtig für das Auftreten, oder nicht Auftreten eines Stereotype Threat Effekts sein. Unter Gruppenidentifikation ist laut Schmader (2002) das Ausmaß, in dem eine Person die Zugehörigkeit zu einer Gruppe als wichtigen Teil der Selbstidentität betrachtet, zu verstehen. So wurde von ihm in einer Studie festgestellt, dass bei Frauen, die eine hohe Identifikation mit dem

eigenen Geschlecht zeigten, bei einer Aktivierung des mathematischen Geschlechtsstereotyps, eine Reduktion der erreichten Mathematikleistung einherging. Bei Frauen welche eine niedrige Identifikation mit ihrer Geschlechtsgruppe zeigten, traten die Stereotype Threat Effekte nicht auf.

Ein weiterer Faktor, der die Wirkung von Stereotype Threat beeinflusst, ist die Identifikation mit der Domäne. Martiny und Götz (2011, S. 163) bezeichnen die Identifikation mit der Domäne als „...das Ausmaß zu welchem eine Person einer bestimmten Domäne Wert beimisst.“ Des weiteren schreiben Martiny und Götz (2011, S. 163) das je wichtiger einem die Domäne ist, „desto wahrscheinlicher werden selbstwertdienliche Strategien angewandt“ und desto wahrscheinlicher werden negative Emotionen mit der Aktivierung negativer Stereotype ausgelöst.

Gerade wenn Menschen eine Domäne besonders wichtig ist, sie sich sehr für sie interessieren, können Stereotype Threat Effekte zum Tragen kommen. Osborne und Walker (2006) begannen 1999 eine Längsschnittstudie, in der sie einige Schüler über zwei Jahre begleiteten. Bei Eintritt in die neunte Klasse wurde den Schülern ein Test vorgelegt, welcher messen soll, wie wichtig den Testpersonen selbst ihre akademischen Leistungen sind. Ein weiterer Fragebogen sollte die Facetten der Identifikation, wie Zugehörigkeitsgefühl und welchen Stellenwert die Schule einnahm, messen. Zum Ende jedes Jahres überprüften die Autoren, ob die Schüler noch die Schule besuchten, ihre GPA (grade point average) sowie Tage die sie fehlten und wie oft sie wegen unangemessenem Verhalten zum Direktor mussten. Die Ergebnisse dieser Studie zeigten, dass Schüler mit höheren GPA seltener abwesend waren oder wegen unangemessenen Verhaltens vorsprechen mussten als jene mit niedrigen GPAs. Die Identifikation mit akademischen Bereichen zeigte ein unterschiedliches Bild für die hell- und dunkelhäutigen Personen. In Übereinstimmung mit den Hypothesen von Steele (1997) zeigte sich, dass akademisch stigmatisierte Personen, welche eine hohe akademische Identifikation aufweisen, eine höhere Wahrscheinlichkeit haben die Schule abzubrechen, weil eine hohe Identifikation die Schule zu einem sehr aversiven Ort macht.

Steele (1997) argumentiert, dass man um Erfolg in der Schule zu haben, eine Identifikation mit Schulleistungen als einen Teil der Selbstdefinition ansehen müsste. Ein positives Selbstgefühl sei also auch zu einem gewissen Maß von guten Leistungen abhängig. Damit diese Identifikation stattfinden kann, müsste man sich gute Chancen in dieser Domäne ausmalen können. Daher können für den Aufbau einer Domänenidentifikation Interessen, Fähigkeiten, Ressourcen, Zugehörigkeits- und Akzeptanzgefühle und eine erlebte Wertschätzung an bzw. in einer bestimmten Domäne als Voraussetzungen angesehen werden (Steele, 1997). So könnte eine Frau, welche eine positive Identifikation mit einer mathematischen Domäne anstrebt, sich mit den niedrigen Erwartungen von Lehrern und den gesellschaftlichen Genderrollen bezüglich Mathematik konfrontiert sehen. Diese Hindernisse würden es ihr schwer machen, eine positive Perspektive in diesen Bereichen zu sehen, was wiederum die Identifikation mit dieser Domäne erschweren würde. Daher wird angenommen, dass Personen, welche über eine hohe Identifikation in einem Bereich verfügen, sich in diesem auch schon beweisen mussten. Good, Aronson und Harder (2008) untersuchten dies bei Studentinnen an der Harvard Universität die als Hauptfach Mathematik studierten. In Folge der Wahl des Studiums gingen die Autoren davon aus, dass die Teilnehmerinnen ein positives Selbstkonzept bezüglich ihrer mathematischen Fähigkeiten haben würden. In ihrer Studie fanden Good et al. (2008) heraus, dass sogar Frauen, welche sich am oberen Ende des Fähigkeitsniveaus in Mathematik befinden, anfällig für die Auswirkungen von Stereotype Threat Effekten waren. Frauen, welche den ST Bedingungen ausgesetzt waren, schnitten in einem speziell für ihren Kurs konstruierten Infinitesimalrechnungstest schlechter ab, als Männer und Frauen in der Kontrollgruppe. In der Kontrollbedingung, in der Gendergleichheit bei dem Test angegeben wurde, schnitten weibliche Teilnehmer sogar besser ab als männliche. Doch nicht alle Autoren sehen die Domänenidentifikation als einen wichtigen Faktor für die Leistungsreduktion unter Stereotype Threat an. In einer Metaanalyse befassten sich Nguyen und Ryan (2008) mit weiblichen Personen und ihren mathematischen Leistungen unter dem Einfluss von Stereotype Threat. Es zeigte sich, dass niedrig-identifizierte Frauen eine geringere Anfälligkeit für den Stereotype Threat Effekt aufwiesen als hoch-

identifizierte Frauen. Am anfälligsten für die Effekte von Stereotype Threat zeigten sich in ihrer Metanalyse moderat mathematik-identifizierte Frauen. Die Autoren argumentieren, dass die hoch-identifizierten Frauen möglicherweise mit Reaktanz auf die präsentierten negativen Stereotype reagieren.

Die Schwierigkeit der Aufgaben hat sich als weiterer wichtiger Moderator für das Auftreten von Stereotype Threat und den daraus resultierenden Konsequenzen für die Leistungen erwiesen. Nach Keller (2007) zeigt sich der negative Effekt der Stereotypaktivierung, nur bei schwierigen Aufgaben. In seiner Studie untersuchte er den Einfluss von Domänenidentifikation und Aufgaben Schwierigkeit auf das Auftreten bzw. die Stärke des Stereotype Threat Effekts bei Realschülern. Es wurden eine Reihe verschiedener Mathematikaufgaben aus unterschiedlichen Quellen mit verschiedenen Lösungswahrscheinlichkeiten verwendet. Die Resultate der Studie zeigten, dass weibliche Teilnehmer mit einer hohen Mathematikidentifikation Leistungseinbußen unter Stereotype Threat zeigten. Außerdem zeigte sich, dass bei schwierigen Aufgaben Personen mit einer niedrigen Mathematikidentifikation höhere Leistungen unter Stereotype Threat erbrachten, als unter der Gendergleichheit Bedingung. Ein genau entgegengesetzter Trend zeigte sich für Personen mit einer hohen Mathematikidentifikation. Diese Interaktion konnte bei einfachen Aufgaben nicht nachgewiesen werden. O'Brien und Crandall (2003) konnten einen ähnlichen Effekt in ihren Studien nachweisen. Frauen schnitten unter Stereotype Threat besser in einem einfachen Mathematiktest und schlechter in einem schwierigen Mathematiktest ab als Frauen, welche nicht unter dem Einfluss von Stereotype Threat standen. Diese Effekte wirkten sich vor allem auf die Genauigkeit mit der gearbeitet wurde bzw. die allgemeine Leistung aus, jedoch nicht auf die Anzahl der Aufgaben, die bearbeitet wurden. Diese Studien zeigen Übereinstimmung mit dem Gedanken von Zajonc (1965; zitiert nach O'Brien & Crandall, 2003, S. 787), dass Erregung (arousal) die Testleistung bei einfachen Aufgaben erhöht und bei schwierigen Aufgaben beeinträchtigt. Zusammengefasst kann man sagen, da man vor allem bei schwierigen Aufgaben alle vorhandenen kognitiven Ressourcen braucht und Stereotype Threat die Kapazitäten des Arbeitsgedächtnisses beeinträchtigt (Schmader & Johns, 2003), kann man vor allem bei solchen

Aufgaben mit einer Leistungseinbuße mit hoher Wahrscheinlichkeit rechnen. Alle beschriebenen Rahmenbedingungen spielen eine wichtige Rolle beim Auftreten des Stereotype Threat Effekts.

1.6.3 Integriertes Prozessmodell der Stereotype Threat Effekte auf die Leistung

Schmader, Forbs und Johns (2008) entwickelten ein Integriertes Prozessmodell der Effekte von Stereotype Threat auf die Leistung. Zur Erstellung des Modells beschäftigten sich die Autoren ausführlich mit Stereotype Threat Effekten im Kontext von Stress, Erregung, Vigilanz, Arbeitsgedächtnis sowie Selbstregulation. Die Autoren geben an, dass Stereotype Threat die Leistung mittels drei distinkten, jedoch nicht zusammenhängenden, Mechanismen stört. Als erstes nennen die Autoren physiologische Stressreaktionen, welche sich direkt auf die präfrontale Verarbeitung auswirken. Des Weiteren wird eine Tendenz aktiv, die eigene Leistung zu überwachen und im Dienste der Selbstregulation Bemühungen gestartet, negative Gedanken und Gefühle zu unterdrücken. Diese Mechanismen machen es uns schwer, bei kognitiven und sozialen Aufgaben, gute Leistungen zu erbringen, da sie einen Großteil unserer exekutiven Ressourcen aufbrauchen. Das Modell von Schmader et al. (2008) steht in Übereinstimmung mit der Theorie der sozialen Identität. Die Theorie der Sozialen Identität geht davon aus, dass Menschen danach streben sowohl eine positive personale Identität, als auch eine positive soziale Identität zu erlangen und beizubehalten (Tajfel & Turner, 1979/1986; zitiert nach Martiny und Götz, 2011, S. 156). Steele, Spencer und Aronson (2002) verstehen Stereotype Threat auch als eine Aktivierung eines negativen Stereotyps, welche eine Bedrohung der sozialen Identität auslöst. „Der Wert der eigenen Person, wie auch der Wert der Gruppe, wird dabei durch soziale Vergleiche bestimmt, (Theorie des sozialen Vergleichs; Festinger, 1954; zitiert nach Martiny und Götz, 2011, S. 156). Menschen sind also dazu geneigt sich selbst, wie auch die Gruppe, der sie angehören, mit anderen Personen und Gruppen zu vergleichen. Dabei hofft man natürlich auf ein positives Ergebnis, um eine positive soziale, wie auch personale Identität zu schützen und beizubehalten (Martiny & Götz, 2011). Da bei einer Bedrohung der sozialen und personalen Identität (wie es bei Stereotype Threat Situationen der Falls ist) die Motivation

ausgelöst wird, diese zu erhalten, sind nicht mehr alle kognitiven Ressourcen verfügbar, um sich mit der vor einem liegenden Aufgabe zu beschäftigen (vgl. Martiny und Götz, 2011).

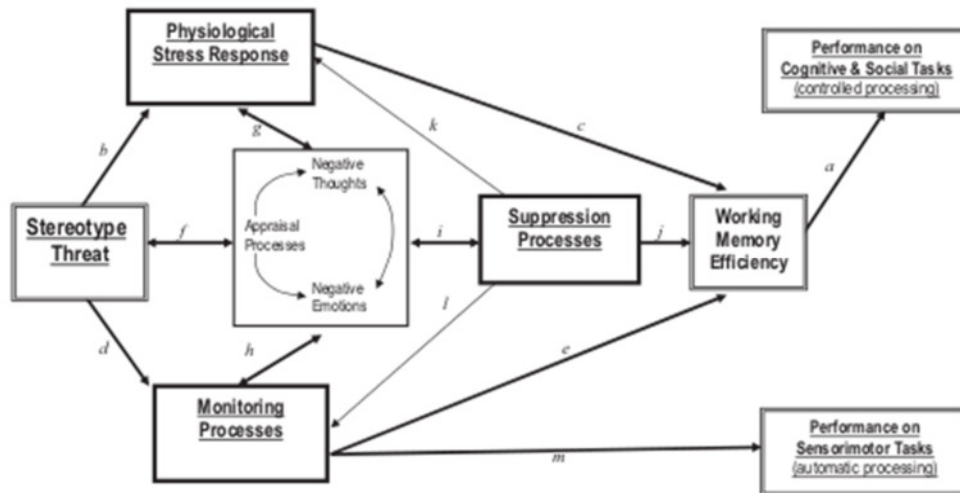


Abbildung 1: Integriertes Prozessmodell der Wirkung von Stereotype Threat auf Leistung
Von Schmader, Forbs und Johns (2008, S. 337)

Abbildung 1 zeigt eine vereinfachte Darstellung des integrierten Prozessmodells der Wirkung von Stereotype Threat. Die drei wichtigen Prozesse werden hier nochmals dargestellt und ihre Wirkung auf das Arbeitsgedächtnis, sowie die Leistung bei kognitiven und sozialen Aufgaben aufgezeigt. Schmader et al. (2008) identifizierten das Arbeitsgedächtnis als exekutive Ressource, welche mit einer effizienten Leistung in Verbindung steht. Diese Ressource muss für eine Reihe von kognitiven und sozialen Aufgaben Informationen koordinieren, während sie gleichzeitig die Interferenz von ablenkenden, oder störenden Informationen zu verhindern versucht. Diese Verbindung wird in der Abbildung 1 durch Pfad a dargestellt. Um es für sich selbst zu verdeutlichen, was die Situation, in der man sich befindet für einen selbst und/oder die Gruppe zu der man gehört, impliziert, werden zwei Prozesse genannt, welche Schmader et al. (2008) als „Physiologische Stress Reaktion“ (dargestellt durch Pfad b in Abbildung 1) in Kombination mit der „Überwachung von Reizen“ (dargestellt durch Pfad d in Abbildung 1) definieren. Diese Prozesse haben einen Einfluss auf die Effizienz des Arbeitsgedächtnisses (dargestellt durch Pfad c und e).

Es wurde angenommen, dass durch die erhöhte Überwachung in Kombination mit erhöhtem physiologischen Stress zusammen mit einem Prime, welcher zu einem Zustand der kognitiven Unausgeglichenheit führt (erzeugt durch Stereotype Threat), Personen dazu verleitet werden könnten, ihre Erfahrungen auf eine verzerrte Art zu betrachten, welches wiederum zu negativen Gedanken und Gefühlen (Pfad *f*, *g* und *h* in Abbildung 1) führen würde. Die Theorie von Schmader et al.'s (2008) besagt, dass Personen, welche unter Stereotype Threat stehen, motiviert sind ein negatives Stereotype nicht zu bestätigen, indem sie besonders gut abschneiden. Sie würden in dem Bestreben besonders gut abzuschneiden, Anstrengungen unternehmen stereotypische und angstbezogene Gedanken, welche mit ihrem Ziel inkonsistent sind, zu unterdrücken (dargestellt durch Pfad *i*). Innerhalb dieses Prozesssets gibt es drei primäre Gründe, aus denen es zu einer Reduktion der Leistung kommen könnte. Diese sind: erstens eine direkte physiologische Beeinträchtigung der präfrontalen Verarbeitung, erzeugt durch die Aktivierung der Hypothalamus-Hypophysen- Nebennieren-Achse (Pfad *c*), zweitens erhöhte Vigilanz gegenüber endogenen oder exogenen Reizen, um sich selbst innerhalb der Situation zu bewerten (Pfad *e*) und drittens aktive Anstrengungen, um stereotype Gedanken und unangenehme Gefühle zu unterdrücken (Pfad *j*). Schmader et al. (2008) beteuern, dass ihr Modell ebenfalls als eine Erklärung für die Effekte von Stereotype Threat auf automatische Prozesse herangezogen werden kann (Pfad *m*).

1.6.4 Mediatoren von Stereotype Threat

Schmader und Johns (2003) konnten eine Reduktion der Kapazität des Arbeitsgedächtnisses bei Latinos und Frauen in ihren Studien nachweisen. Außerdem konnten sie diese Reduktion in der Kapazität des Arbeitsgedächtnisses als ein Mediator für das Auftreten der mathematischen Leistungsminderung von Frauen unter Stereotype Threat identifizieren. Croizet, Després, Gauzins, Huguet, Leyens und Méot (2004) konnten in ihren Studien sogar einen Einfluss der Aktivierung negativer Stereotypen auf die Herzratenvariabilität (welche die Autoren als physiologischen Indikator für die mentale Belastung ansahen) nachweisen. Es konnte ebenfalls mittels Selbsteinschätzung nachgewiesen werden das Arousal (Erregung) als Mediator für den Zusammenhang zwischen

der Aktivierung eines negativen Stereotyps und reduzierter Leistung in Frage kommt (Martiny, Roth & Deaux, 2010; zitiert nach Martiny & Götz, 2011, S. 159).

Negative Emotionen kommen bereits in dem integrierten Prozessmodel von Schmader et al. (2008) vor. Sie wirken sich bei Aktivierung eines negativen Stereotyps auf die verfügbaren Ressourcen des Arbeitsgedächtnisses aus und beeinträchtigen darüber auch die Leistungen. Auch Steele (1997) spricht im Zusammenhang mit Stereotype Threat von der Angst, einem negativen Stereotyp zu entsprechen. Keller und Dauenheimer (2003) konnten Frustration (dejection) als Mediator für den Zusammenhang zwischen dem Auftreten der Leistungseinbußen bei (mittel bis wenig mit der relevanten Domäne identifizierten) Personen, welche zu der von dem negativen Stereotyp betroffenen Gruppe gehörten und der Aktivierung des negativen Stereotyps identifizieren.

Da durch die Aktivierung von negativen Stereotypen negative Gefühle angeregt werden (Steele, 1997; Keller und Dauenheimer 2003), welches sich auch in physiologischer Erregung bemerkbar macht (Croizet et al., 2004) liegt die Annahme, dass dies auch negative, oder aufgabenirrelevante Gedanken erzeugen könnte, nicht allzu fern (Martiny & Götz, 2011). Steele und Aronson (1995) konnten in einem ihrer Experimente zeigen, dass die stereotypisierte Personengruppe (in diesem Fall waren dies dunkelhäutige Personen), welche glaubte einen diagnostischen Test zu machen, eine signifikant höhere Aktivierung von Stereotypen über dunkelhäutige Personen hatte. Zudem zeigten diese Personen eine höhere Aktivierung der Befürchtungen, bezüglich ihrer Fähigkeiten und sie vermieden mit einer höheren Tendenz rassenbezogene stereotype Präferenzen. Außerdem zeigte sich bei den Zielpersonen, dass sie bereits im Vorhinein Ausreden für ihre Leistungen machten und es widerstrebte ihnen deutlich mehr als den anderen, nicht stereotypisierten Personen, dass ihre Rassenidentität mit ihren Leistungen in Verbindung gebracht werden könnten. All dies sind Hinweise für die Vielzahl an möglichen aufgabenirrelevanten Informationen welche Zielpersonen, die unter dem Einfluss der Aktivierung negativer Stereotype stehen, möglicherweise von ihren Aufgaben ablenken und somit ihre Leistungen beeinträchtigen. Auch im direkten Zusammenhang mit unserem Forschungsschwerpunkt konnten Forscher bereits Hinweise auf den

mediierenden Effekt von negativen Emotionen, auf die mathematische Leistungen von Frauen unter Stereotype Threat, finden. Ein Beispiel dafür stellen die Ergebnisse von Cadinu, Maass, Rosabianca und Kiesner (2005) dar. Mittels einer Gedankenauflistungstechnik dokumentierten sie die Gedanken von Frauen, bei denen ein negativer Stereotyp aktiviert wurde und verglichen sie mit denen einer Kontrollgruppe. Frauen bei denen ein negativer Stereotyp aktiviert worden war, berichteten von einer höheren Anzahl negativer Gedanken, welche in Verbindung mit Mathematik und dem zu absolvierenden Test standen, als jene Frauen in der Kontrollbedingung. Darüber hinaus zeigte sich ebenfalls bei den Frauen unter Stereotype Threat ein deutlicher Leistungsabfall, (vor allem während der zweiten Hälfte des Tests), welcher durch die negativen Gedanken mediert wurde. Um nicht den Rahmen unserer Untersuchung zu sprengen, haben wir darauf verzichtet die Gedanken unserer Teilnehmer aufzuzeichnen. Wir wollten jedoch mit Hilfe der Aufnahme der Arbeitshaltung (wie schwierig die Untersuchung empfunden wurde usw.) zumindest einen Teil dieser Gedanken möglicherweise nachvollziehen können.

Als Erweiterung zu Schmader et al.'s (2008) Modell fügten Martiny und Götz (2011) noch den Mediator Motivation hinzu. Wobei der Einfluss von motivationalen Variablen sowohl beim Erleben der durch den Stereotyp empfunden Bedrohung, als auch ihr Einfluss auf die Leistungserwartungen der Zielpersonen gesehen wird. Cadinu, Maass, Frigerio, Impagliazzo und Latinotti (2003) untersuchten in ihrer Studie wie sich die Erwartungshaltung und Leistung in einem schweren Mathematiktest durch vorangegangene Manipulation verändert. Den verschiedenen Versuchsgruppen wurde mitgeteilt, dass Frauen entweder besser, schlechter, oder gleich gut wie Männer, bei dem vor ihnen liegenden Test, abschnitten. Zudem wurden noch Identifikation mit der Domäne (Mathematik), sowie die Kontrollüberzeugung erhoben. Es zeigte sich, dass die ausschließlich weiblichen Teilnehmer in der Bedingung mit der positiven Information die besten Ergebnisse erzielten und jene mit der negativen die schlechtesten. Dieser Trend zeigte sich jedoch nur bei den Teilnehmerinnen mit einer hohen Domänenidentifikation. Personen, welche die negative Information erhalten hatten und logisch-mathematische Fähigkeiten als wichtig ansahen, zeigten auch die

geringsten Leistungserwartungen und die schlechtesten Leistungen in dem Mathematiktest. Die Erwartungshaltung konnte von den Autoren als ein partieller Mediator des Effekts von Stereotype Threat auf Leistung identifiziert werden.

Harkins (2006) argumentierte, dass Personen, welche Stereotype Threat unterliegen, eine Motivation haben in den vor ihnen liegenden Aufgaben gut abzuschneiden. Dies potenziert wiederum präpotente (reflexartige) Reaktionen, aber auch das Bestreben aus, dieser Tendenz entgegen zu wirken, sobald die Teilnehmer bemerkten, dass ihre Reaktion falsch war, was die richtige Reaktion gewesen wäre und sie die Möglichkeit hatten, diese wahrzunehmen. Der „Mere Effort“ Ansatz von Harkins (2006) bezüglich Stereotype Threat nimmt an, dass Stereotype Threat ähnlich wirkt, wie eine potentielle Evaluation und dies wiederum sollte Personen dazu animieren, gut in einer Aufgabe abzuschneiden. Jamieson und Harkins (2007) untersuchten diese Theorien mittels Inhibition Aufgaben. Konsistent mit den Erwartungen des Mere Effort Ansatzes, zeigte sich, dass Personen in der Stereotype Threat Bedingung schlechter abschnitten, als Personen der Kontrollgruppe. Wenn den Teilnehmern allerdings die Möglichkeit gegeben wurde, ihre Antworten zu korrigieren, schnitten Personen unter Stereotype Threat sogar besser ab, als Personen der Kontrollbedingung. Jamieson und Harkins (2007) erklären dies dadurch, dass bei Personen unter Stereotype Threat die präpotenten Reaktionen zwar erhöht wurden, aber sie auch besonders motiviert waren gut abzuschneiden und daher eher die Möglichkeit zur Korrektur wahrnahmen.

1.7 Stereotype Threat im Bezug auf Gender und Mathematikleistung

In Zusammenhang mit dem gängigen Stereotyp „Frauen sind schlechter als Männer in mathematischen Bereichen“ wurde eine Vielzahl an Untersuchungen durchgeführt, von denen bereits einige Erwähnung fanden. Eine der am häufigsten genannten Studien im Zusammenhang mit Stereotype Threat, ist jene von Steele (1997). Er konnte feststellen, dass sowohl Frauen, welche an dieses Stereotyp glaubten, als auch jene, die dies nicht taten, eine Verschlechterung der Leistung unter ST erlitten. Nach Steele (1997) müssen allerdings ganz bestimmte

Bedingungen erfüllt sein, damit ST auftritt. Die negative Wirkung des Stereotyps zeigt sich, wenn die betroffene Person sich über ihren Erfolg in einer Umgebung (in unserem Fall wäre das die Schule) definiert und mit dem Stereotyp konfrontiert wird. Quinn und Spencer (2001) zeigten in ihre Untersuchungen unter anderem, dass ST Effekte bei Frauen, die mathematische Textaufgaben zu lösen hatten, am ehesten auftraten. Weiblichen Versuchspersonen fiel es besonders schwer, geeignete Problemlösungsstrategien bei dieser Art von Aufgabe zu entwickeln. Wie bereits in Abschnitt 1.4.2 erwähnt, spielen auch die Identifikation mit der Domäne, sowie die Aufgabenschwierigkeit eine große Rolle. Stereotype Threat Effekte traten vor allem bei schwierigen Aufgaben und bei Personen mit einer hohen Mathematikidentifikation auf (O'Brien & Crandall, 2003; Keller, 2007).

In einer Studie von Johnson et al. (2012) werden ST und SL Effekte anhand einer Stichprobe unter amerikanischen Studenten einander gegenüber gestellt. Die genauen Ergebnisse dieser Studie sowie die für unsere Untersuchung daraus relevanten Resultate und die daraus abgeleiteten Hypothesen werden im folgenden Kapitel im Detail besprochen.

1.8 Stereotype Lift

Durch Stereotypenpriming kann nicht nur eine Leistungsminderung erreicht werden. Unter besonderen Umständen lässt sich mit Hilfe des Stereotypenprimings auch eine Leistungssteigerung erreichen. Dieses Phänomen wird Stereotype Lift (SL) genannt. Stereotype Lift kann als „...the performance boost caused by the awareness that an outgroup is negatively stereotyped.“ beschrieben werden (Walton & Cohen 2003, S. 456). In einer Meta-Analyse fanden Walton und Cohen (2003) heraus, dass es kaum einen Unterschied auf den Stereotype Lift Effekt hatte, ob die Studien beim Priming explizit auf einen Stereotyp hinwiesen (z. B. das der verwendete Test Unterschiede zwischen den Leistungen der verschiedenen Gruppen zeigte), oder nur Angaben eine bestimmte Fähigkeit diagnostisch zu untersuchen. Die Verbindung zwischen einem negativen Stereotyp der Gruppe, welcher man selbst nicht angehört, (also der Außengruppe oder outgroup) und den evaluativen Tests ist so stark, dass es keinen zusätzlichen Nutzen hat diesen explizit zu erwähnen. Walton und Cohen (2003) fanden

Hinweise darauf, dass in Studien bei denen entweder die Validität des Stereotyps oder dessen Relevanz für den Test widerlegt werden sollte, ein stärkerer Lift Effekt zu messen war, als bei Studien die dies nicht taten. Außerdem scheint Stereotype Lift wie auch Stereotype Threat (Steele, 1997) einen größeren Effekt zu haben, wenn es den getesteten Personen wichtig erscheint gut abzuschneiden (Walton & Cohen, 2003). Stereotype Lift und *Stereotype Susceptibility* (bzw. Stereotype Anfälligkeit) sind ähnliche Phänomene, jedoch mit einem wichtigen Unterschied. Stereotype Susceptibility ist eine Leistungssteigerung, welche durch die Aktivierung von positiven Stereotypen über die eigene Gruppe verursacht wird (Shih, Pittinsky, & Ambady, 1999). Stereotype Lift Effekte werden nach Walton und Cohen (2003) allerdings durch die Aktivierung eines negativen Stereotyps über eine andere Gruppe verursacht.

Zwei Mediatoren für das Auftreten von Stereotype Lift wurden von Chatard, Selimbegovic, Konan und Mugny (2008) identifiziert. Ähnlich wie bereits bei Stereotype Threat (Schmader, Johns & Barquissau, 2004), konnten *Stereotype Endorsement* (Befürwortung von Stereotypen) und Vorurteile im Hinblick auf SL als Mediatoren identifiziert werden. Das untersuchte Stereotyp, mit welchem sich Chatard et al. (2008) beschäftigen, beinhaltete Aussagen über unterschiedliche intellektuelle Fähigkeiten von inländischen Kindern (Franzosen) und afrikanischen Immigranten. Stereotype Endorsement wurde mittels zweier Aussagen, bezüglich der intellektuellen Fähigkeiten von französischen und afrikanischen Immigranten, erfasst. Das Ausmaß der Vorurteile wurde mit Hilfe von vier Begriffen mit verschiedenen Aussagen zum Thema Immigranten erfasst. Eine Aussage war beispielsweise: "I think that the French government should be stricter on African immigration" (Chatard et al., 2008, S.1422). Teilnehmer wurden gebeten, ihre Zustimmung bezüglich der Aussagen auf einer Skala von eins bis sieben anzugeben. Teilnehmer, welche hohe Werte in Stereotype Endorsement hatten und jene mit hohen Werten in der Vorurteilsbefragung, schnitten signifikant besser in der Bedingung ab, in der das vorherrschende Stereotyp bewusst gemacht wurde, als in jener, in der nur gesagt wurde, dass der Test individuelle Unterschiede messen würde. Bei Personen mit niedrigen Werten in genannten Mediatoren war dies nicht der Fall.

Ein weiterer Faktor, welcher das Auftreten von Stereotype Lift beeinflusst, ist die Aktivierung des social self (sozialen Selbst) in stereotyprelevanten (diagnostischen) Testsituationen. Marx und Stapel (2006) aktivierten das soziale Selbst, indem sie Probanden baten in einem Absatz alle „wir“, „uns“ usw. einzukreisen. In einem zwei (Geschlecht der VP) x fünf (Soziales Selbst aktiv „wir“ + Diagnostischer Test, „Ich“ Prime + diagnostischer Test, Genderangabe + Diagnostischer Test, Diagnostischer Test, Kontrollgruppe) Versuchsdesign verglichen sie verschiedene Manipulationsmethoden miteinander. Als Test wurde die Mathematiksektion des Graduate Record Exam (GRE) verwendet. Marx und Stapel (2006) fanden, dass es nicht reichte nur die Testsituation als diagnostisch für bestimmte Fähigkeiten zu beschreiben, da dies nicht zu starken Lift Effekten führte außer, dass das soziale Selbst aktiviert wurde. Auf weibliche Teilnehmer ihrer Studie hatte die Aktivierung des sozialen selbst nicht den gleichen Effekt, wie auf männliche, sie schnitten in allen Versuchsbedingungen schlechter ab als jene, in der Kontrollgruppe.

Stereotype Lift Effekte beschränken sich jedoch nicht nur auf kognitive Bereiche wie z.B., Leistungstests. Chalabaev, Stone, Sarrazin und Croizet (2008) konnten einen solchen Effekt auch bei der Ausführung von motorischen Fähigkeiten nachweisen. Die Autoren sagten den Teilnehmern explizit, dass ein Geschlecht besser in einer Balanceaufgabe sei, als das andere und ließen sie dann je nach Versuchsbedingung noch ein Video von einem Mann, oder einer Frau ansehen, welche/r sehr gut in der Balanceaufgabe abschnitt. Die Kontrollgruppe bekam keine geschlechtsspezifischen Informationen. In der Stereotype Lift Bedingung bekamen beispielsweise weibliche VP die Information, Männer würden schlecht bei solchen Aufgaben abschneiden und sahen anschließend ein Video einer Frau, welche sehr gut in der Aufgabe abschnitt. Alle Testpersonen mussten einen Manipulationscheck sowie einen Fragebogen zum kompetitiven Beunruhigung nach dem Priming, jedoch vor dem Ausführen der Balanceaufgabe, ausfüllen. Alle Gruppen wurden nach dem Priming gebeten, die zuvor im Video gesehene Aufgabe selbst auszuführen. Chalabaev et al. (2008) konnten einen Stereotype Lift Effekt sowohl bei weiblichen, als auch bei männlichen Personen nachweisen. Bei den männlichen Probanden konnten sie feststellen, dass dieser Effekt mit

hoher Selbstzuversicht und Aufgabeninvolvierung einherging. Die Autoren schlussfolgern, dass der explizite Vergleich nach unten mit einer Aufwertung der nicht vom Stereotyp betroffenen Gruppe (in diesem Fall war das die Gruppe, welcher die Probanden selbst angehörten) die Selbstzuversicht sowie Aufgabeninvolvierung verbesserte und dass diese Variablen zum beobachteten Leistungssteigerung beigetragen haben. Daher können Selbstzuversicht und Aufgabeninvolvierung, als Mediatoren für das Auftreten von Stereotype Lift gesehen werden.

1.9 Hypothesen

Durch die in der Literatur angeführten Inhalte, welche sich mit Stereotype Threat und Lift Effekten befasste, wurden acht Haupthypothesen und vier Nebenhypothesen formuliert, welchen wir nachgehen wollten. Um möglichst gut erklären zu können, in welche Richtung und in welcher Stärke der Stereotype Threat Effekt in unserer Zielpopulation auftreten sollte, wurde entschieden möglichst viele Fragen innerhalb der demographischen Daten abzuhandeln.

1.9.1 H1: Gibt es einen signifikanten Unterschied in der Basisfähigkeit Mathematik zwischen Buben und Mädchen?

In den USA werden jährlich flächendeckende Schulleistungstests im Rahmen des NCLB (No Child is left behind) Acts bzw. Gesetzes durchgeführt. Das Ziel dieser Tests ist es, dass Eltern und Lehrern mehr Verantwortung für die Leistungen ihrer Kinder übernehmen, da Schulen die bei diesen Tests schlecht abschneiden, sanktioniert werden. Ein weiterer gewünschter Effekt dieses Gesetzes wäre es Kindern, welche aus verschiedenen Gründen als benachteiligt gelten, die gleichen Bildungschancen zu geben, wie anderen Schülern (Lee, 2014). Hyde et al. (2008) ließen sich für eine Metaanalyse in Bezug auf Genderunterschiede in mathematischen Bereichen, die statistischen Daten der NCLB von zehn repräsentativen Staaten geben. Hyde et al. (2008) konnten in dieser Metaanalyse der Mathematikergebnisse aus zehn Staaten keinen signifikanten Geschlechtsunterschied in mathematischen Bereichen feststellen. Ähnliche Ergebnisse bei den Metastudien von Scafidi und Bui (2010), sowie einer weiteren Metaanalyse von Lindberg, Hyde, Petersen und Linn (2010). Diese stehen diese

Ergebnisse im Widerspruch zu den Ergebnissen der PISA Studie, welche durchaus Geschlechtsunterschiede im Mathematischen Bereich gefunden hatte. Möglicherweise sind die angeführten Studien nur für die USA repräsentativ. Da sich widersprüchliche Aussagen bezüglich der Mathematischen Fähigkeiten von Mädchen und Burschen in der Literatur finden, wird diese Hypothese ungerichtet formuliert.

1.9.2 H2: Weibliche Versuchspersonen, welche einen Stereotypnonkonformen (SL) Artikel erhalten, erbringen bessere oder gleich gute Leistungen bei dem Mathematiktest nach dem Priming, wie im Mathematiktest vor dem Priming

Weibliche Teilnehmer, die den Artikel zu lesen bekamen, welcher zum Ausdruck brachte, dass Frauen besser in mathematischen Bereichen abschneiden als Männer, sollten dem Stereotyp Lift Effekt unterliegen. Die Leistungen der Teilnehmerinnen unter Stereotyp Lift sollen sich signifikant verbessern, oder zumindest gleich bleiben. Johnson et al. (2012) fanden keinen Hinweis auf einen Stereotype Lift Effekt bei amerikanischen Studentinnen, jedoch schnitten weibliche Personen unter der Stereotype Lift genauso gut ab, wie die Kontrollgruppe und besser, als unter Stereotype Threat. Da Walton und Cohen (2003) jedoch postulierten, dass Individuen unter Stereotype Lift besser abschneiden, als unter anderen Bedingungen, formulierten wir unsere Hypothesen entsprechend.

1.9.3 H3: Weibliche Versuchsteilnehmer, welche einen Stereotypkonformen Artikel (Stereotype Threat) erhalten, bringen im Mathematiktest nach dem Priming schlechtere Leistungen als im Mathematiktest vor dem Priming

Diese Hypothese wurde nach dem Stereotype Threat Effekt formuliert. Dieser besagt, dass bestimmte Stereotype über die Leistung einer bestimmten Gruppe in der Gesellschaft existieren und das Individuen, welche dieser Gruppe angehören, sich so verhalten, dass sie diese negativen Stereotype bestätigen (Steele, 1997). Ein solcher Stereotyp besteht bezüglich der Leistung von Frauen im

mathematischen Bereich. Übereinstimmend mit den Ergebnissen von Spencer, Steele und Quinn (1999) sowie Johnson et al. (2012) wurde erwartet, dass die weiblichen Probanden, welche der Stereotype Threat Bedingung zugeteilt werden würden, schlechtere Leistungen erbringen als unter Stereotyp Lift und im Baseline Mathematiktest. Auch diese Hypothese wurde gerichtet formuliert

1.9.4 H4: Weibliche Teilnehmer schneiden im Mathematiktest nach dem Priming unter Stereotype Lift besser ab als unter Stereotype Threat

Aus diesen Überlegungen ergab sich auch, dass weibliche Personen unter Stereotype Lift besser abschneiden würden, als unter Stereotype Threat. Unsere Hypothese wurde auch hier gerichtet formuliert.

1.9.5 H5: Männliche Versuchsteilnehmer, welche der Stereotype Lift Bedingung ausgesetzt werden, bringen genauso gute Leistungen im Mathematiktest vor, wie im Mathematiktest nach dem Priming

Konform mit den Ergebnissen von Johnson et al. (2012) erwarteten wir keine Leistungsveränderung der männlichen Teilnehmern, welche der Stereotype Lift Bedingung ausgesetzt werden. Zwischen dem Baseline Mathematiktest und dem Mathematiktest, welcher nach dem Priming vorgegeben wurde, sollte daher keine Leistungsveränderung stattfinden. Unsere Hypothese wird gerichtet formuliert:

1.9.6 H6: Männliche Versuchsteilnehmer, welche der Stereotype Threat Bedingung ausgesetzt werden, bringen im Mathematiktest nach dem Priming bessere Leistungen, als im Mathematiktest vor dem Priming

Den Ergebnissen von Johnson et al. (2012) zufolge schnitten Männer, welcher unter Stereotype Threat gesetzt wurden, besser ab als jene unter Stereotype Lift.

1.9.7 H7: Männliche Versuchsteilnehmer in der Stereotype Threat Bedingung schneiden im Mathematiktest nach dem Priming gleich gut ab wie männliche Versuchsteilnehmer in der Stereotype Lift Bedingung.

Ebenfalls konform mit den Ergebnissen von Johnson et al. (2012) wurde erwartet, dass männliche Versuchsteilnehmer, welche der Stereotype Threat Bedingung ausgesetzt werden, gleich gut abschneiden wie männliche Personen in der Stereotype Lift Bedingung bezüglich der erreichten Punktezahl im Mathematiktest, welcher nach dem Priming vorgegeben wird.

1.9.8 H8: Männliche Versuchsteilnehmer schneiden im Mathematiktest nach dem Priming unter ST besser ab als Frauen unter ST

Da Männer in der Stereotype Lift Bedingung, laut den Ergebnissen von Johnson et al. (2012), gleich gut abschneiden wie unter Stereotype Threat, während Frauen sich unter Stereotype Threat verschlechtern, kann man daraus folgern, dass Männer unter Stereotype Threat besser abschneiden werden als Frauen, welche der Stereotype Threat Bedingung ausgesetzt werden. Entsprechend wurde unsere Hypothese formuliert.

1.10 Weitere Fragestellungen

Nach dem Studium der relevanten Literatur wurden noch einige Nebenhypothesen formuliert.

1.10.1 Hat Interesse an Mathematik einen Einfluss auf Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde?

Schiefele, Krapp und Schreyer (1993) konnten in einer Metaanalyse, des Zusammenhangs zwischen der Ausprägung fachspezifischer Interessen und schulischer Leistung, eine Durchschnittskorrelation von $r = 0.30$ finden. Die Autoren fanden eine höhere Interessen-Leistungskorrelation für männliche Personen, als für weibliche Personen. Ob ein hohes Interesse an Mathematik

allerdings vor dem schädlichen Einfluss von Stereotype Threat schützen kann, bleibt abzuwarten. So zeigte sich, dass auch Personen mit hohem Interesse an dem Fach, dem leistungsschädigenden Einfluss von negativen Stereotypen unterliegen (Schmader, Johns & Barquissau, 2004).

1.10.2 Hat die persönliche Wichtigkeit von Mathematik einen Effekt auf die Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde?

Die Variable der persönlichen Wichtigkeit sollte einerseits als Domänenidentifikation verwendet werden, andererseits sollte der Frage nachgegangen werden, ob möglicherweise eine größere empfundene Wichtigkeit des Faches auch einen Effekt auf die Leistungen im Mathematiktest vor dem Priming sowie auf die Stärke des Stereotype Threat Effekts hat.

1.10.3 Beeinflusst das Priming die Selbsteinschätzung?

Nachdem den Teilnehmern eine bestimmte Information gegeben wurde, welche sich auf ihr Geschlecht (oder das andere Geschlecht) bezog, erwarteten wir, wie bereits in einigen Studien nachgewiesen worden ist, einen entsprechenden Effekt auf die Einschätzung ihrer eigenen Leistungen.

1.10.4 Beeinflusst das Priming die Fremdeinschätzung?

Um zu ermitteln ob sich die Einschätzung von weiblichen und männlichen Teilnehmern bezüglich mathematischen Leistungen „anderer Mädchen und Burschen ihres Alters“ nach dem Priming ändert, erfragen wir nach den jeweiligen Mathematiktests was die Teilnehmer annehmen wie viele Aufgaben andere Mädchen und Burschen ihres Alters richtig beantwortet haben.

1.10.5 Wenn Teilnehmer eine sehr geringe Motivation aufweisen, hat das Priming keinen Einfluss auf die Mathematiktestwerte der Teilnehmer

Da unsere Teilnehmer keine Sonderpunkte oder etwas Ähnliches für die Teilnahme an der Studie bekamen, befürchteten wir eine geringe Motivation für die Teilnahme und somit auch dass z.B. der Artikel, durch den das Priming

erfolgen sollte, von den Teilnehmern nicht gelesen werden würde. Wir wollten daher sichergehen, dass ein gewisser Grad an Motivation vorhanden sein würde, damit das Priming einen Effekt haben könnte.

1.10.6 Hat die Primingmethode einen Einfluss auf die Stärke des Stereotype Threat, oder Stereotype Lift Effekt?

Wie bereits besprochen gibt es verschiedene Methoden mit welchen sich ein Stereotype Threat Effekt hervorrufen lässt. An dieser Stelle ist die Frage von Interesse, ob sich diese Methoden in der Stärke des hervorgerufenen Effekts (d.h. der Leistungsminderung) unterscheiden. Von Interesse ist ebenfalls die Frage, ob es Unterschiede in der Wirkung von direktem Priming, wie es Appel et al. (2011) verwendeten und einem subtileren Ansatz, wie er von Lang (2004) verwendet wurde, gibt.

2 Methode

2.1 Untersuchungsdesign

In dem vorangegangenen Kapitel wurde ausführlich über Priming und Stereotype Threat berichtet. In dieser Studie sollten, statt wie üblich mit einer Kontrollgruppe, die Leistungen der Schüler vor und nach einem Priming verglichen werden. Die Entscheidung bezüglich der zu untersuchenden Stichprobe fiel auf Schulkinder der 9. Schulstufe, da sich in diesem Altersbereich bereits Stereotype Threat (Huguet, P. & Régner, 2008) und Lift Effekte (Chatard et al., 2008) nachweisen ließen. Außerdem wurde angenommen, dass Jugendliche in diesem Alter bereits über gute PC Kenntnisse verfügen und so keine Schwierigkeiten mit der ausgewählten Datenerhebungsmethode haben würden. Ein Programm für Onlineumfragen wurde zur Datenerhebung ausgewählt, da dies die Testauswertung erleichtert und mögliche Fehler, die bei der Übertragung von Papier-Bleistift-Tests zu SPSS auftreten können, ausschließt. Das genaue Vorgehen zur Datenerhebung wird in Kapitel 2.4 beschrieben.

2.1.1 Durchführung der Untersuchung

Das Experiment wurde im Zeitraum von November 2013 bis Februar 2014, als kontrollierte Onlinestudie an vier Gymnasien in Wiener Neustadt durchgeführt. Es wurden Schüler und Schülerinnen der 9ten Schulstufe untersucht. Die Testung fand gruppenweise statt, wobei jeder Schüler bei der Bearbeitung der Aufgaben einen PC zur Verfügung hatte. Die Testung fand an zwei ca. eine Woche auseinander liegenden Terminen während der regulären Unterrichtszeit statt.

Die Instruktionen sowie der Ablauf waren so ausgelegt, dass diese selbsterklärend und auch ohne weitere Erklärung, oder Hilfe auszuführen waren. Eine bis zwei Testleiterinnen standen den Teilnehmern während der Bearbeitung für Nachfragen und für die Behebung eventueller technischer Probleme zur Verfügung. Die Versuchspersonen wurden explizit dazu angehalten, alleine und selbstständig die Aufgaben zu bearbeiten. Ihnen wurde die Möglichkeit gegeben bei der Bearbeitung der mathematischen Aufgaben die jeweiligen Rechenwege aufzuschreiben. Den Schülerinnen wurde es frei gestellt, ihren eigenen

Taschenrechner, oder den im Programm integrierten zu verwenden. Die Lehrkräfte, welche Ihre Stunden zur Verfügung stellten, waren nicht bei jeder Klasse anwesend. Jene, welche der Testung beiwohnten, griffen jedoch nicht in die Testung ein.

Die Testung wurde mittels einer freien Online Umfrage-Applikation namens LimeSurvey durchgeführt. Bereitgestellt wurde die Untersuchung auf einem privaten Server.

2.2 Untersuchungsinstrumente

2.2.1 CFT 20 R Subtest Matrizen

Der CFT 20 R (Weis, 2006) ist ein vielseitig einsetzbarer Intelligenztest, welcher das allgemeine intellektuelle Niveau (Grundintelligenz) im Sinne der "General Fluid Ability" nach Cattell erfasst. Sein Einsatzbereich erstreckt sich bei Kindern und Jugendlichen auf einen Altersbereich von 8.5 bis 19 Jahre. Für diese Studie wurde nur die Kurzform der Subtest Matrizen 1 verwendet. Vom CFT 20 R wurden die Bestandteile sieben bis 16 verwendet, da die ersten sechs Aufgaben des Tests von den meisten Personen leicht gelöst werden. Zur ersten Testung wurden die geraden Aufgaben (z.B. Item sechs, Item acht usw.) und beim zweiten Testzeitpunkt die ungeraden Aufgaben zur Bearbeitung aufgegeben. Die Teilnehmer hatten bei jedem Termin eine Minute und vierzig Sekunden Zeit, um die vorgegebenen Aufgaben zu lösen. Der Zeitrahmen für die Bearbeitung der Aufgaben wurde an die neue Vorgabeform angepasst. Nach dem Ausfüllen der demographischen Daten folgte eine Beispielaufgabe aus dem Matrizen 1 Test des CFT 20 R und eine Erklärung wie die folgenden Aufgaben zu bearbeiten seien (siehe Abbildung 2). In der Erklärung stand:

„Nun folgen einige Denkaufgaben. Hierbei sollst du jeweils das Muster finden, dass am besten in das leere Kästchen passt. Versuche sie so schnell wie möglich der Reihe nach zu lösen.

Sobald du auf „Weiter“ klickst wird ein Timer gestartet. Du hast 1:40 Zeit die folgenden Aufgaben zu bearbeiten.

Bearbeite nun zuerst die Beispielaufgabe!“

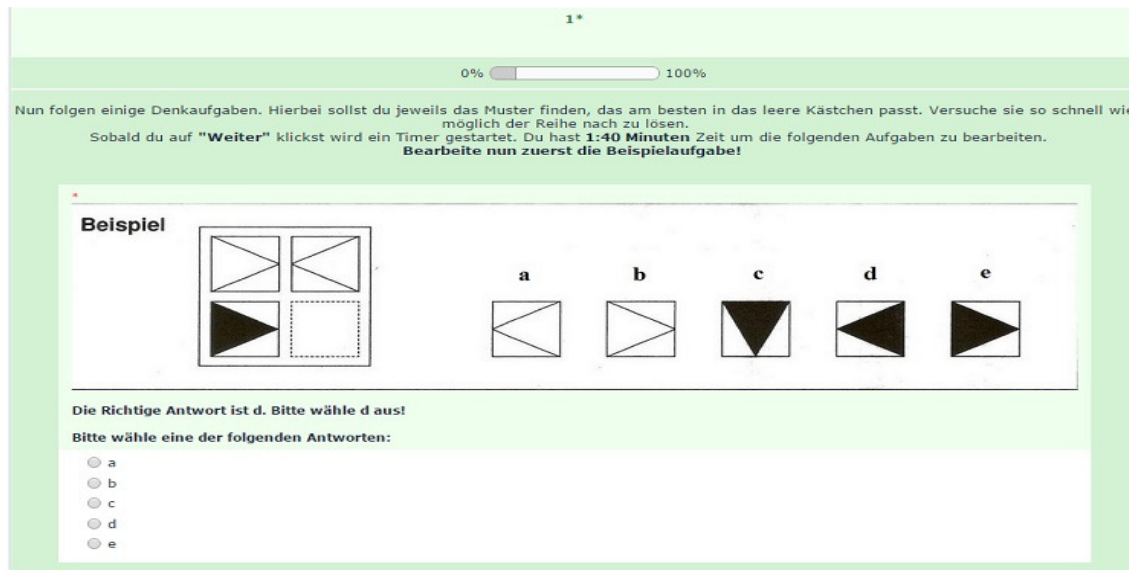


Abbildung 2: Ein Screenshot der ersten Beispielaufgabe des CFT 20 R Matrizen Tests mit dazugehöriger Instruktion

Nur wenn die Schüler der Instruktion folgten und die Beispielaufgabe richtig lösten (die richtige Antwort stand in der Instruktion) konnten sie die fünf darauffolgenden Testaufgaben bearbeiten. Das Zeitlimit wurde mittels eines eingebauten Timers des Programms LimeSurvey kontrolliert.. Wenn die Zeit zur Bearbeitung abgelaufen war konnten die Teilnehmer keine der Antwortmöglichkeiten mehr anklicken und mussten auf die nächste Seite wechseln. Nach der Bearbeitung der CFT 20 R Aufgaben sollten die Schüler einschätzen, wie viele der fünf Aufgaben sie richtig beantwortet hatten und zu schätzen, wie viele der Aufgaben andere Mädchen und Jungen ihres Alters richtig beantwortet hatten. Durch das Programm war die Eingabe auf die Zahlen von 0 bis 5 beschränkt.

2.2.2 Baseline Mathematiktest und Parallel Versionen

Der *Baseline Mathematiktest* bestand aus acht Aufgaben, welche aus verschiedenen Bereichen der Mathematik stammen. Der Baseline Mathematiktest ist jener Mathematiktest, welcher vor dem jeweiligen Priming vorgegeben wurde und wird in den folgenden Abschnitten auch aus *Pre-Priming Mathematiktest* bezeichnet. Die Aufgaben wurden an ähnliche Aufgaben aus verschiedenen Mathematikübungsbüchern der vierten und fünften Schulstufe angelehnt. Die Teilnehmer hatten insgesamt 26 Minuten für die acht Aufgaben des Baseline Mathematiktests zur Verfügung. Da die Aufgaben des Baseline Mathematiktests

auf zwei Termine aufgeteilt waren, hatten die Schüler 13 Minuten Zeit für die Bearbeitung von vier Aufgaben pro Termin.

Der Mathematiktest welcher nach dem Priming vorgegeben wurde, hatte acht Aufgaben, welche am Stück auf einer Seite vorgegeben wurden. Diese waren, wie auch jene des Pre-Priming Mathematiktests der Schwierigkeit nach angeordnet. Das Zeitlimit für den Mathematiktest, welcher nach dem Priming stattfand betrug 25 Minuten. Um Verwirrung zu vermeiden wird der Mathematiktest, welcher nach dem Priming vorgegeben wurde auch als *Post-Priming* Mathematiktest bezeichnet.

Der Instruktionstext, welcher auf einer einzelnen Seite vor dem Beginn der Mathematiktests angezeigt wurde, war folgender:

„Es folgen nun einige Rechenaufgaben. Bitte versuche, sie der Reihe nach zu bearbeiten. Du hast für alle Aufgaben nur eine gewisse Zeit - wie viel Zeit noch übrig ist, wird dir am Bildschirm angezeigt.

Wahrscheinlich kannst du nicht alle Aufgaben in der Zeit lösen. Bemühe dich bitte trotzdem, so rasch und sorgfältig wie möglich zu arbeiten.

Du kannst den Taschenrechner am Bildschirm benutzen, wenn du willst. Wenn du Zwischenschritte aufschreiben möchtest, benutze bitte das Blatt, dass du bekommen hast.

Viel Erfolg“

Der Instruktionstext war bei allen Terminen und Mathematiktests identisch.

Es wurden mehrere Mathematiklehrer von verschiedenen Gymnasien gebeten, sich die Aufgaben anzusehen und ihrer Schwierigkeit nach zu sortieren. Die Lehrer legten uns nahe, nicht zu viele Aufgaben zu wählen, da sie selbst auch nicht mehr als zehn Aufgaben in ihren Mathematikschultests vorgeben. Sie geben ihnen für gewöhnlich eine ganze Schulstunde Zeit, um diese Aufgaben nach bestem Wissen zu beantworten. Uns war es ein Ziel, ein ähnliches Schwierigkeitsniveau wie in den Schularbeiten zu erreichen. Wir wählten eine gerade Anzahl an Aufgaben, damit eine gleichmäßige Aufteilung der Aufgaben für den Baseline Test möglich war.

Die Zahlen der Aufgaben wurden so gewählt, dass die Aufgaben auch ohne Taschenrechner gelöst werden konnten, bzw. die Antworten nicht zu viele Kommastellen aufwiesen. Die Aufgaben stammen aus den Bereichen der Geometrie und der Prozentrechnung. Zudem waren Bewegungsaufgaben, Mischungsaufgaben und Gleichungssysteme dabei. Zu jeder Aufgabe des Pre-Priming Mathematiktests gab es eine Parallelaufgabe im Post-Priming Mathematiktest. Parallelaufgabe bedeutet in diesem Sinne, dass die Aufgabe einen ähnlichen Satzbau und Lösungsweg aufwies wie ihr Pendant. Es wurden Textaufgaben verwendet, da gilt, dass die Bearbeitung dieser Aufgaben sehr anfällig für Priming bzw. Stereotype Threat Effekte ist. (Quinn und Spencer, 2001) Wir haben uns bemüht geschlechtsneutrale Themen für die Textaufgaben zu verwenden. Es wurden insgesamt zwei Parallelversionen des Baseline (Pre-Priming) Mathematiktests erstellt. Einem ging ein Priming mittels eines Artikels voran, dies ist auch jener Mathematiktest mit dem sich diese Diplomarbeit befasst. Dem zweiten ging ein anderes Priming mittels Adjektivpaaren voran. Näheres zu dieser Primingmethode und der zweiten Parallelversion des Mathematiktests findet sich in der Diplomarbeit meiner Kollegin Regina Reischer.

2.2.3 Priming mittels Artikel

Der verwendete Artikel ist eine abgekürzte und vereinfachte Version des Artikels, welcher von Appel et al. (2011) für den Artikel „Stereotype Threat impedes ability building: Effects on test preparation among women in science and technology“ genutzt wurde. Der Artikel wurde mit Einverständnis des Autors, welcher ihn übersetzte, verwendet. Für die Testung in den Schulen stand nur eine limitierte Zeit zur Verfügung, daher wurde der Artikel verkürzt, um den Teilnehmern mehr Zeit für die Bearbeitung der mathematischen Aufgaben zur Verfügung zu stellen.

Die Schüler erhielten die Instruktion den Artikel genau zu lesen, um im Anschluss Fragen bezüglich des Inhalts des Artikels, beantworten zu können. Der genaue Instruktionstext war:

„Lies bitte den folgenden Text aufmerksam durch und versuche Dir soviel wie möglich zu merken. Es werden Dir im Anschluss Fragen dazu gestellt.“

Den Schülern wurde je nach Versuchsbedingung, ein Text mit der Information gegeben, dass Mädchen bzw. Jungen besser in mathematischen Bereichen wären und dass dies durch eine große und ausführliche Studie bestätigt worden sei. Im Anschluss an den Text folgten acht Mathematikaufgaben. Nach der Bearbeitung der acht Mathematikaufgaben, welche dem Priming folgten, wurden die Teilnehmer gebeten verschiedene Fragen, zu den Inhalten des Artikels, zu beantworten. Diese Fragen waren zum Teil jene welche Appel et al. (2011) in seiner Testung verwendete. Die Fragen wurden einerseits gestellt um das Lesen des Artikels zu rechtfertigen und andererseits, um zu kontrollieren ob die Teilnehmer den Artikel auch wirklich gelesen hatten. Die wichtigste Frage war jene, nach der Grundaussage des Artikels. Die Versuchspersonen konnten in Aussagen formulierte Sätze entweder mit „richtig“ oder „falsch“ ankreuzen. Außerdem wurden von mir noch zwei weitere Fragen erstellt, welche im Anhang zu finden sind. Mit diesen Fragen hoffte ich zu ergründen, ob die Teilnehmer möglicherweise mehr oder weniger den Effekten des Primings unterlagen, weil sie den Artikel für sehr glaubwürdig hielten, oder ob sie mit dem Stereotype übereinstimmten das der jeweiliger Artikel suggeriert hatte.

2.2.4 Fragebogen zum Arbeitsstil

Dieser Fragebogen wurde von Lang (2004) und Meidlinger (2003) für Ihre Diplomarbeiten entwickelt, um die Auswirkungen des Primings auf die Arbeitsmotivation sowie den Arbeitsstil zu erheben. Dieser Fragebogen bestand aus sieben Aussagen (Abbildung 3). Bei diesen Aussagen handelte es sich um entgegengesetzte Adjektivpaare, welche anhand des semantischen Differenzials formatiert waren. Anders als beim Semantischen Differenzial üblich wurde eine gerade Anzahl von Antwortmöglichkeiten gewählt um ein „durchklicken“ bzw. eine Tendenz zur Mitte zu minimieren.

Beantworte bitte zum Schluss noch ein paar kurze Fragen darüber, wie es dir beim Bearbeiten des gesamten Fragebogens ergangen ist.

* Bei der Bearbeitung der Fragen war ich...					
☐ konzentriert	☐ .	☐ .	☐ .	☐ .	☐ unkonzentriert
* .					
☐ schnell	☐ .	☐ .	☐ .	☐ .	☐ langsam
* .					
☐ nicht interessiert	☐ .	☐ .	☐ .	☐ .	☐ interessiert
* .					
☐ motiviert	☐ .	☐ .	☐ .	☐ .	☐ unmotiviert
* .					
☐ überfordert	☐ .	☐ .	☐ .	☐ .	☐ unterfordert
* .					
☐ wenig bemüht	☐ .	☐ .	☐ .	☐ .	☐ sehr bemüht
* Die Rechenaufgaben waren für mich...					
☐ einfach	☐ .	☐ .	☐ .	☐ .	☐ schwierig

Abbildung 3: Ein Screenshot des Fragebogens zum Arbeitsstil

2.3 Vorgehen

Die Untersuchung wurde über den Verlauf von zwei ungefähr eine Woche auseinanderliegenden Terminen durchgeführt. Der genaue Ablauf der Untersuchung, sowie die Zuteilung der Versuchspersonen zu den Versuchsbedingungen wird detailliert in Abschnitten 2.4.1 und Abschnitt 2.4.2 dargestellt.

2.3.1 Ablauf des ersten Testtermins

Zu Beginn des ersten Termins wurden die Schüler aus ihren Klassen abgeholt. Jeder Schüler wurde gebeten sich an einen PC zu setzen, diesen einzuschalten und sich einzuloggen. Es folgte eine kurze mündliche Begrüßung, in der die Schüler über den ungefähren Ablauf der Untersuchung aufgeklärt wurden. Jeder Schüler erhielt dann einen Papierbogen (Anhang C) aus drei doppelseitig beschrifteten DIN A4 Seiten, auf denen sie ihren persönlichen Code notieren und ihre Rechenwege aufschreiben konnten. Die Beschriftung bestand aus einem Feld für das Datum, einen Feld für den persönlichen Code (samt Erklärung, wie dieser vierstelligen Code zu erstellen ist) und Feldern mit Nummerierung für die

Mathematikaufgaben. Ein kleiner Zettel mit der URL der jeweiligen Testversion war in der oberen Ecke angebracht.

Die Schüler wurden gebeten, die tinyURL³, welche sich an ihrem Papierbogen befand, in einen Browser ihrer Wahl einzugeben und den Instruktionen am Bildschirm zu folgen. Bei Fragen konnten die Schülerinnen sich melden und die Testleiter ansprechen.

Die Instruktion des Startbildschirms war folgende:

„Herzlich willkommen!

Vielen Dank, dass Du dich entschlossen hast, bei unserer Studie mit zu machen. Du leistest damit einen wertvollen Beitrag zu unserer Arbeit.

Die Teilnahme an dieser Studie ist freiwillig.

Versuche so sorgfältig und zügig wie möglich zu arbeiten und sämtliche Fragen ehrlich zu beantworten.“

Es gab insgesamt acht verschiedene tinyURL Codes, welche jeweils zu einer bestimmten Version unserer Testung führten. Vier der URLs führten zu dem ersten Teil unserer Testung, welche beim ersten Testzeitpunkt vorgegeben wurde und vier enthielten den zweiten Teil unserer Testung. Die acht Versionen der Testung unterschieden sich im Zeitpunkt der Vorgabe (nur bei diesen Versionen wurden die demographischen Daten abgefragt) in der verwendeten Primingmethode (P) (es kamen zwei verschiedene Primingmethoden zum Einsatz) und hinsichtlich des Geschlechts, auf das geprimt wurde. Außerdem wurden die Matrizenaufgaben des CFT 20 R sowie die Aufgaben des Mathematik Baseline Tests auf die zwei Termine verteilt. Es wurde versucht, gleichwertige Mathematikaufgaben für die insgesamt drei Versionen des Tests zu erstellen. Jeder Teilnehmer sollte zum Schluss der

³ Eine „tinyURL“ ist eine gekürzte Form eines langen Links, welcher mit online verfügbaren URL Shortenern generiert werden kann.

Untersuchung einen kompletten Baselinemathematiktest (B1, B2), CFT Matrizentest (CFT_{T1}/CFT_{T2}), Mathematiktest Version M1 und Mathematiktest Version M2 absolviert haben.

Zum besseren Verständnis des Testablaufs wird dieser mittels Abbildung 4 bildlich dargestellt.

Erster Termin



Zweiter Termin



CODE	Persönlicher Code	B1	Baseline Mathematiktest Teil 1	^W P _M	Priming Artikel / BEM Weiblich/Männlich
D	Demographische Daten	B2	Baseline Mathematiktest Teil2	M1	Mathematiktest Version 1 Artikel
CFT _{T1}	CFT Aufgaben Teil 1	NB _{B1}	Nachbefragung B1	M2	Mathematiktest Version 2 BEM
CFT _{T2}	CFT Aufgaben Teil 2	NB _{B2}	Nachbefragung B2	NB _{M1}	Nachbefragung M1
G	Fragen zum Artikel	S	Selbstbeurteilung des BEM	NB _{M2}	Nachbefragung M2
AH	Arbeitshaltungen				

Abbildung 4: Eine schematisierte Darstellung des ersten und zweiten Teils des Untersuchungsablaufes

Bei dem ersten Termin wurde nach der Erstellung eines persönlichen Codes die Eingabe der demographischen Daten von den Schülern verlangt. Im Rahmen der demographischen Daten (D) wurden folgende Daten erhoben: das Geschlecht, Geburtsjahr, Geburtsmonat, Zuhause am meisten gesprochene Sprache, im letzten Jahr besuchte Schule/Klasse, das persönliche Interesse an Mathematik, Deutsch, Englisch, für wie wichtig sie die Fächer Mathematik, Englisch, Deutsch für ihre Zukunft ansahen, welche Noten sie im letzten Zeugnis in Mathematik/Deutsch/Englisch bekommen hatten, für wie gut sie ihre Fähigkeiten in Mathematik/Deutsch/Englisch hielten und für wie gut sie die Fähigkeiten von anderen Burschen bzw. Mädchen ihres Alters einschätzten.

Nach der Eingabe der demographischen Daten (D) der Teilnehmer folgten je nach Version, entweder die geraden, oder ungeraden ersten fünf Aufgaben des CFT20 R (CFT_{T1}/CFT_{T2}) und des Baseline Mathematiktests, sowie jeweils eine Nachbefragung (NB_{B1} / NB_{B2}), um herauszufinden, wie viele Aufgaben die Teilnehmer nach eigener Einschätzung korrekt hatten lösen können. Für die Bearbeitung der ersten Hälfte Baseline Mathematiktests (B1), welcher aus 4 Aufgaben bestand, hatten die Teilnehmer 13 Minuten Zeit. Die Zeit wurde mittels eines im Limesurvey verankerten Timers begrenzt und nach dessen Ablauf konnten keine Zahlen mehr in die dafür vorgesehenen Felder eingetragen werden. Auf die Nachbefragung folgte das Priming (P). Es kamen zwei Priming-Methoden zum Einsatz. Eine Version war ein Priming mit Hilfe eines Artikels, welcher das Thema Mathematik und Geschlecht beinhaltete. Die Kinder bekamen entweder die Information, dass im Rahmen einer neuen Studie herausgekommen sei, dass Mädchen besser in mathematischen Bereichen abschneiden (W), oder das Jungen besser in mathematischen Bereichen abschneiden (M). Die Kinder wurden angewiesen den Artikel genau zu lesen, da im Anschluss Fragen zu dessen Inhalt gestellt werden würden. Die zweite Primingversion bestand aus Adjektivlisten, entnommen aus dem Bem Sex Roles Inventory (BSRI; Bem, 1981), bei dem sich die Versuchspersonen entweder in eine weibliche (W), oder männliche Person (M) hinein versetzten sollen. Statt der Beantwortung von Fragen, wurden die Teilnehmer im Anschluss an den Mathematiktest (M2) gebeten, den Bem Sex Roles Inventory nochmals aus ihrer Sicht auszufüllen (S). Diese Diplomarbeit befasst sich mit der ersten Version des Primings, einem stereotypaktivierenden Artikel. Nach dem Priming folgte ein weiterer Mathematiktest, wobei der M1 Mathematiktest immer mit der Artikel-Priming-Methode gekoppelt wurde und der M2 Mathematiktest immer mit der anderen Methode. M1 wie auch M2 stellen Parallelversionen des Baseline Mathematiktests dar. Auf den Mathematiktest M1 (bzw. M2) folgte immer eine Nachbefragung (NB_{M1}/NB_{M2}) darüber, wie viele Aufgaben die Kinder glaubten selbst richtig gelöst zu haben und was sie dachten, wie viele Aufgaben Mädchen und Jungen ihres Alters richtig gelöst hatten. Auf die Nachbefragung des Mathematiktest M1 folgten einige Aussagen, welche sich auf den Artikel bezogen (G). Diese Aussagen konnten als „richtig“, oder

„falsch“ bewertet werden. Außerdem wurde noch nach der Glaubwürdigkeit des Artikels gefragt und ob die Teilnehmer mit den Aussagen des Artikels übereinstimmten. Die letzte Aufgabe diente der Ermittlung der „Arbeitshaltungen“ (AH). Die Teilnehmer sollten auf einer Skala von 1 bis 6 einschätzen wie motiviert sie waren, wie schwierig die Aufgaben für sie waren, wie konzentriert sie waren, wie bemüht sie waren, ob sie eher über/ oder unterfordert waren und wie schnell sie bei der Bearbeitung der Aufgaben waren.

Die Papierbögen mit den Rechenschritten (Anhang C) wurden eingesammelt und die Teilnehmer am Ende des ersten Termins darauf hingewiesen, dass es einen zweiten Testzeitpunkt geben würde.

2.3.2 Ablauf des zweiten Testzeitpunktes

Vor Beginn des zweiten Testzeitpunktes wurden die persönlichen Codes der Teilnehmer aus LimeSurvey ausgelesen. Um eine gleichmäßige Verteilung der Geschlechter in den verschiedenen Priming-Bedingungen zu haben, wurden die Teilnehmer je nach Geschlecht und Priming, welches sie beim ersten Termin bekommen hatten, Gruppen zugeordnet. So bekam beispielsweise eine weibliche Versuchsperson, welche beim ersten Termin den Artikel mit der Aussage „Frauen sind besser in mathematischen Bereichen“ beim zweiten Termin das Priming, bei welchem sie sich in eine männliche Person ihres Alters hineinversetzen sollte. Keine der Versuchspersonen durchlief zweimal die gleiche Priming Variante. Weibliche Versuchspersonen konnten jedoch durchaus beim ersten Termin den Artikel mit der Aussage „Männer sind besser in mathematischen Bereichen“ bekommen und im zweiten Termin die Aufgabe bekommen, sich in eine männliche Person ihres Alters hineinzusetzen. Es ergaben sich insgesamt 32 verschiedene Kombinationsmöglichkeiten der Versuchsbedingungen. Eine andere mögliche Kombination wäre eine weibliche Teilnehmerin, welche beim ersten Termin mit dem Artikel mit der Aussage „Frauen sind besser in mathematischen Bereichen“ und beim Zweiten das indirekte Priming, mit der Bitte sich in eine weibliche Person ihres Alters hinein zu versetzen, erfahren. Eine weibliche Teilnehmerin konnte auch beim ersten Termin mit dem Artikel der Aussage „Männer sind besser in mathematischen Bereichen“ konfrontiert werden

und beim zweiten Termin, das indirekte Priming mit der Bitte sich in eine weibliche Person ihres Alters hinein zu versetzen, bekommen. Für männliche Teilnehmer galt diese Regel analog. Außerdem wurde der einen Hälfte der Teilnehmer zuerst der Artikel vorgegeben und der anderen Hälfte das indirekte Priming. Abbildung 5 zeigt einige der möglichen Kombinationen des Versuchsablaufs. „VP“ steht in Abbildung 5 für Versuchsperson.

	Testzeitpunkt 1	Testzeitpunkt 2
Männliche VP	Priming A Weiblich	Priming B Weiblich
Weibliche VP	Priming A Weiblich	Priming B Weiblich
Männliche VP	Priming A Weiblich	Priming B Männlich
Weibliche VP	Priming A Weiblich	Priming B Männlich
Männliche VP	Priming B Weiblich	Priming A Weiblich
Weibliche VP	Priming B Weiblich	Priming A Weiblich
Männliche VP	Priming B Männlich	Priming A Weiblich
Weibliche VP	Priming B Männlich	Priming A Weiblich

Abbildung 5: Mögliche Kombinationen der Primingversionen und des Geschlechts der Testperson über die beiden Termine verteilt

Diese Diplomarbeit befasst sich allerdings nur mit den Ergebnissen des Primings mittels des Artikels (Priming A).

Die Persönlichen Codes der Teilnehmer wurden an eine Tafel geschrieben und die Kinder wurden gebeten sich entsprechend ihres Codes in verschiedenen Räumen, oder Bereichen des PC Raumes zu platzieren. Ihnen wurde dann wie beim ersten Termin, ein Papierbogen aus drei doppelseitig beschrifteten DIN A4 Seiten gegeben, auf denen sie ihren persönlichen Code notieren und ihre Rechenwege aufschreiben konnten. Auf einem kleinen, an diesen Papierbogen angehefteten, Zettel stand wieder eine tinyURL, welche die Teilnehmer in einen Browser ihrer Wahl eingeben konnten. Der Link führte die Teilnehmer zum zweiten Teil der Testung. Sie wurden mit einem ähnlichen Text wie beim ersten Termin begrüßt.

Nach der erneuten Eingabe ihres persönlichen Codes, wurde den Teilnehmern wieder eine Beispielaufgabe des Matrizen Test der CFT 20 R vorgegeben. Nach der korrekten Beantwortung der Beispielaufgabe, konnten die Teilnehmer die beim ersten Termin nicht vorgegebenen Teile des CFT 20 R, bearbeiten. Die Teilnehmer hatten für die Bearbeitung der fünf Aufgaben eine Minute und vierzig Sekunden Zeit. Wie beim ersten Termin folgte den CFT 20 R Matrizen Aufgaben eine Nachbefragung.

Im Anschluss an die Nachbefragung zu den CFT 20 R Matrizen folgte der zweite Teil der Aufgaben des Baseline Mathematiktests. Der Zweite Teil des Baseline Mathematiktests bestand aus vier Aufgaben, die ihrer Schwierigkeit nach angeordnet waren. Die Teilnehmer hatten 13 Minuten Zeit für die Bearbeitung der vier Aufgaben. Die Aufgaben wurden alle untereinander auf einer Seite angezeigt und waren mit einer Anzeige versehen, die den Versuchspersonen zeigte, wie viel Zeit sie für die Bearbeitung der Aufgaben noch zur Verfügung hatten. Außerdem war, wie beim ersten Termin, unter jeder Aufgabe ein Taschenrechner eingearbeitet. Den Teilnehmern wurde wieder freigestellt diesen Taschenrechner, oder ihren eigenen zu verwenden. Wie beim ersten Teil folgte auf die Baselinemathematikaufgaben eine Nachbefragung.

Der Nachbefragung (NB B2) folgte das zweite Priming Variante. Wenn die Versuchsperson beim ersten Termin das direkte Priming (Artikel) bekommen hatte, bekam sie beim zweiten Termin das indirekte Priming (Hineinversetzen in eine weibliche oder männliche Person). Kein Teilnehmer bekam zwei Mal die gleiche Priming Variante. Nach dem Priming wurden die Teilnehmer gebeten, die Aufgaben jeweiligen Mathematiktests zu bearbeiten. Für diese Aufgaben hatten sie 25 Minuten Zeit. Diese Aufgaben waren ebenfalls der Schwierigkeit nach angeordnet und den Schülern stand ein Taschenrechner zur Verfügung. Der Parallelversion des Mathematiktests (M2), welcher das indirekte Priming vorangegangen war, folgte wieder eine Nachbefragung (NB M2) bezüglich der Einschätzung ihrer Leistungen und der, der Anderen.

Wenn die Teilnehmer beim zweiten Termin den Artikel gelesen hatten, folgte der Nachbefragung des Mathematiktests (M2) wieder eine Reihe von Fragen, welche

sich auf den Artikel bezogen. Im Anschluss, folgten Fragen über die jeweiligen Arbeitshaltungen. Wenn die Teilnehmer aber das indirekte Priming durchlaufen hatten, wurden sie nach der Bearbeitung der Parallelversion (M1) des Mathematiktests gebeten den Fragebogen nochmals für sich selbst auszufüllen. Am Ende des zweiten Testzeitpunktes wurden die Teilnehmer über die Ziele der Studie aufgeklärt und eventuelle Fragen beantwortet. Die Versuchspersonen wurden gefragt, ob sie das Gefühl hatten, von dem Priming beeinflusst worden zu sein. Nur sehr wenige Teilnehmer gaben an, das Gefühl gehabt zu haben, von dem Priming beeinflusst gewesen zu sein.

2.3.3 LimeSurvey

LimeSurvey ist eine freie online Umfrageapplikation, mit der man auch ohne Programmierkenntnisse online Umfragen erstellen, veröffentlichen und deren Ergebnisse downloaden kann. Sie läuft mit der Programmiersprache PHP und JavaScript. LimeSurvey hat eine benutzerfreundliche Oberfläche, welche es auch dem Laien ermöglicht, eine eigene Onlineumfrage zu erstellen. Das Programm erlaubt einem eine Vielzahl von Einstellungsmöglichkeiten, bezüglich der einzelnen Fragen und deren Codierung. Wir haben uns für dieses Programm aufgrund der vielen Einstellungsmöglichkeiten und der, gegenüber einer manuellen Eingabe der Daten, geringeren Fehlerwahrscheinlichkeit entschieden.

2.3.4 Der persönliche Code

Zu Beginn jeder Testung wurde von den Teilnehmern verlangt einen „persönlichen Code“ zu erstellen und in eine Maske einzugeben. Dieser Code setzte sich zusammen aus dem ersten Buchstaben des Vornamens der Versuchsperson, dem dritten Buchstaben des Mädchennamens der Mutter der Versuchsperson, der letzten Ziffer der Hausnummer in der die Versuchsperson wohnte und des dritten Buchstabens des Geburtsmonats der Versuchsperson. Abbildung 6 zeigt einen Screenshot des Codeeingabefensters.

Abbildung 6: Screenshot des Startbildschirms für die Eingabe des persönlichen Codes

Der in Abbildung 6 dargestellte Screenshot soll verdeutlichen, wie das Eingabefenster aussah. Die Teilnehmer wurden nochmals darüber aufgeklärt, dass unsere Studie anonym durchgeführt wird und dass die Erstellung eines persönlichen Codes notwendig war, um die beiden Fragebögen einander zuzuordnen.

2.4 Beschreibung der Stichprobe

Die Stichprobe bestand aus insgesamt 305 Schülern aus verschiedenen Gymnasien von Wiener Neustadt. Von den 305 Personen wurden 35 ausgeschlossen. Drei dieser Versuchspersonen wurden ausgeschlossen, da sie nach der Testung Angaben „*nicht wirklich mitgemacht*“ zu haben. Eine Klasse wurde ausgeschlossen, da diese in einer ersten Testung weniger Zeit für die Mathematikaufgaben gehabt hatten. Auf deren Arbeitsblättern wurde dies vermerkt und sie wurden anhand ihrer Codes aus dem Datensatz ausgeschlossen. Für die Durchführung der Studie standen uns pro Termin circa 60 bis 70 Minuten zur Verfügung. Die Zeit, die für die Bearbeitung der Mathematikaufgaben zu Verfügung stand, wurde daher mittels eines Timers begrenzt. Da die Teilnehmer der ersten getesteten Klasse viel schneller als erwartet mit den nicht mathematischen Teilen der Testung fertig wurden, gab uns dies die Möglichkeit den folgenden Klassen mehr Zeit für die mathematischen Teile der Testung einzuräumen. Von den verbleibenden 270 Personen waren 139 (51.5%) weiblich und 131 (48.5%) männlich. Das Alter der Teilnehmer reichte von 14 bis 17 Jahre,

der Median betrug 15 Jahre. Als Muttersprache gaben 247 Personen Deutsch, sieben Türkisch, drei Serbisch, drei Englisch, drei Rumänisch, zwei Ungarisch, eine Kroatisch, eine Albanisch, eine Chinesisch und eine Bengalisch an. Als vorherige Schulstufe gaben 107 Personen Gymnasium an, 94 Realgymnasium, 37 Hauptschule und 32 Personen gaben an, im Vorjahr die Mittelschule besucht zu haben.

242 Teilnehmer hatten den Mathematiktest welcher nach dem Priming vorgegeben wurde, komplett ausgefüllt und wurden zur Überprüfung der Güte dieses Mathematiktests herangezogen.

Von den insgesamt 270 Personen, welche nach dem ersten Ausschluss verblieben, waren 225 bei beiden Testzeitpunkten anwesend und konnten für den Vergleich der Priming-Bedingungen herangezogen werden. Von den 225 Personen waren 119 weiblich und 106 männlich, wobei 119 Personen den Artikel mit dem stereotypkonformen Inhalt und 116 Personen den Artikel mit Inhalt erhielten, der mit dem Stereotyp nicht übereinstimmt. Tabelle 1 stellt die Aufteilung der Versuchspersonen bezüglich ihres Geschlechts und der Priming-Bedingung optisch dar.

Tabelle 1: Aufteilung der Studienteilnehmer auf die verschiedenen Versuchsbedingungen.
Hypothesenprüfung

Geschlecht	Priming	N
Weiblich	Stereotyp Nonkonform	61
	Stereotyp Konform	58
Männlich	Stereotyp Nonkonform	55
	Stereotyp Konform	51

2.5 Analyse der Instrumente

2.5.1 CFT 20 R Subtest Matrizen

Vom CFT 20 R wurden die Aufgaben 6 bis 16 verwendet, da die ersten fünf Aufgaben des Tests sehr einfach sind und daher von den meisten Personen gelöst werden. Zur ersten Testung wurden die geraden Aufgaben (z.B. Aufgabe 2, Aufgabe 4 usw.) und beim zweiten Testzeitpunkt die ungeraden Aufgaben gestellt. Beim ersten Testzeitpunkt lösten die Teilnehmer durchschnittlich $M = 3.72$ ($SD =$

1.32) Aufgaben und beim zweiten Testzeitpunkt $M = 3.85$ ($SD = 1.45$). Bei der Betrachtung der Ergebnisse über beide Testzeitpunkte hinweg, lösten die Teilnehmer im Mittel $M = 5.69$ ($SD 2.20$) von 10 Aufgaben. Drei Personen konnten keine einzige Aufgabe korrekt lösen. Der Kolmogorov Smirnov Test zeigt einen Wert von $p = 0.005$ ($N = 241$) für die Variable der erreichten Gesamtpunkte im CFT, dies Ergebnis ist signifikant und daher wird eine Abweichung von der Normalverteilung angenommen.

Die Reliabilitätsanalyse zeigt für alle Aufgaben ähnliche Trennschärfen im Mittelbereich (Tabelle 2) bei einer Gesamtreliabilität von $\alpha = 0.60$. Die geringe Reliabilität könnte sich dadurch erklären, dass den Versuchspersonen die Übung in Folge der sechs nicht verwendeten, leichteren Aufgaben des CFT fehlt und daher die 7 verbleibenden Aufgaben schwieriger wirken.

Tabelle 2: Reliabilitätstabelle für die vorgegebenen CFT20 R Matrizen Aufgaben 6 bis 15. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben

Item-Statistiken CFT 20 R- Matrizen				
	Skalenmittel wenn Item gelöscht	Skalenvarianz wenn Item gelöscht	Item Gesamtkorrelation	Cronbach's Alpha wenn Item gelöscht
CFT 06	4.49	4.541	0.263	0.620
CFT 07	4.65	4.140	0.376	0.595
CFT 08	4.70	4.224	0.304	0.611
CFT 09	4.80	4.259	0.265	0.620
CFT 10	4.77	4.451	0.171	0.641
CFT 11	4.80	3.972	0.418	0.584
CFT 12	4.99	4.468	0.173	0.640
CFT 13	4.75	3.853	0.497	0.565
CFT 14	5.06	4.417	0.224	0.628
CFT 15	5.13	4.259	0.361	0.600

Die Aufgabenschwierigkeiten der einzelnen Aufgaben finden sich in Tabelle 3 wieder. Erwartungsgemäß waren die Aufgaben, welche im letzten Teil des Tests vorkamen, diejenigen mit der geringsten Lösewahrscheinlichkeit. Eine Ausnahme zu dieser Regel ist Aufgabe 13, welches von 67 % der Personen richtig gelöst wurde.

Tabelle 3: Aufgabenschwierigkeiten der CFT 20 R Matrizen Aufgaben. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben

Item	Item-Schwierigkeit	Standard Abweichung
CFT 6	0.85	0.35
CFT 7	0.78	0.41
CFT 8	0.65	0.48
CFT 9	0.62	0.49
CFT 10	0.56	0.50
CFT 11	0.61	0.49
CFT 12	0.37	0.48
CFT 13	0.67	0.47
CFT 14	0.31	0.46
CFT 15	0.25	0.43

2.5.2 Baseline Mathematiktest – Pre Priming

Der verwendete Baseline Mathematiktest, welcher als Maß für die Grundfähigkeiten in Mathematik diente, bestand aus acht mathematischen Textaufgaben. Die verwendeten Aufgaben waren an jene aus Lehrbüchern des Gymnasiums für die 8te und 9te Schulstufe angelehnt. Die Teilnehmer hatten pro Termin 13 Minuten für die Bearbeitung von jeweils vier Aufgaben Zeit. Die Teilnehmer konnten ursprünglich bis zu zwei Punkte pro Aufgabe erreichen. Wenn die Aufgabe korrekt gelöst wurde, erhielt der Teilnehmer zwei Punkte. Wenn die Antwort nicht korrekt war, wurde auf den Arbeitsblättern des Teilnehmers geschaut, ob dieser den richtigen Rechenweg benutzt hatte. Falls dies der Fall war, erhielt der Teilnehmer einen Punkt für die entsprechende Aufgabe. Die maximale Punktzahl sollte dementsprechend sechzehn sein. Da diese Berechnung allerdings keine zusätzliche Information lieferte, wurde im weiteren Verlauf mit einem Punkt für eine korrekte Antwort und null Punkten für eine inkorrekte Antwort gerechnet. Von den 270 Testpersonen, welche nach dem ersten Ausschluss verblieben, hatten 241 einen kompletten Baseline Mathematiktest absolviert, bzw. waren bei beiden Terminen anwesend. Insgesamt wurden 129 weibliche und 112 männliche Teilnehmer für die Berechnungen berücksichtigt. Weibliche Teilnehmer erreichten eine durchschnittliche Punktzahl von $M = 2.9$ Punkten mit einem Median von 3 und männliche Teilnehmer erreichten durchschnittlich $M = 3.22$ Punkte mit einem Median von ebenfalls 3. In Tabelle 4 sind der Gesamtdurchschnitt, wie auch die Standardabweichungen aufgelistet. Der Test erreicht ein Cronbach's Alpha von

$\alpha = 0.41$. Berechnet man die Reliabilität für die Geschlechter getrennt, erreicht der Test bei den weiblichen Teilnehmern ein Cronbach's Alpha von $\alpha = 0.306$ und bei den Männlichen $\alpha = 0.502$.

Tabelle 4: Deskriptive Statistiken des Baseline Mathematiktest

	Weibliche TP	Männliche TP	Gesamt
N	129	112	241
Mittelwert	2.899	3.223	3.050
Standardabweichung	1.286	1.499	1.396
Minimale Punktzahl	0	0	0
Maximale Punktzahl	6	7	7

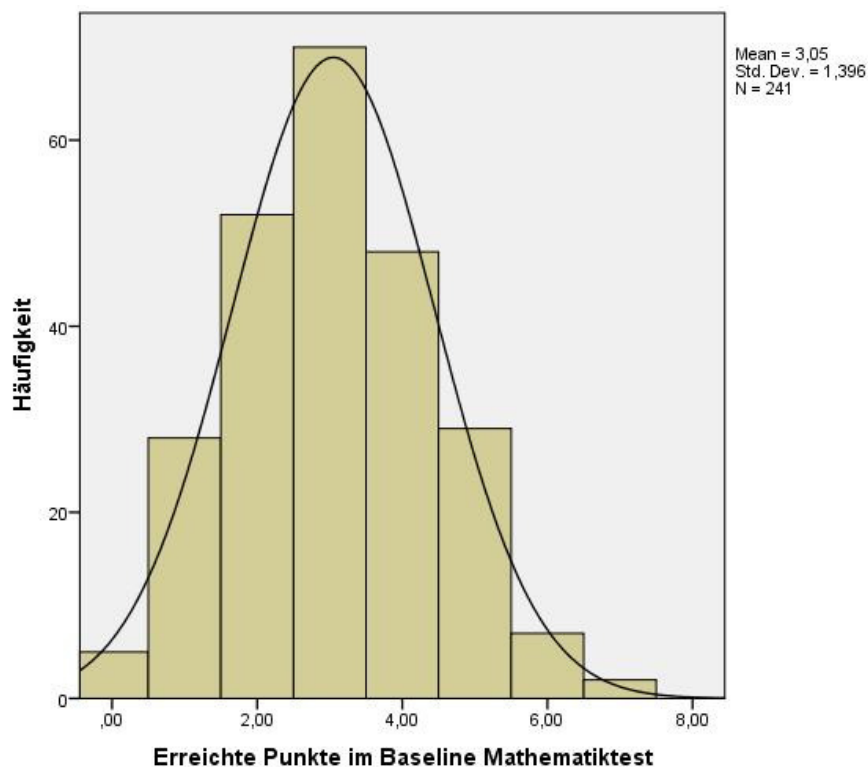


Abbildung 7: Ein Histogramm der Verteilung der von den Teilnehmern im Baseline Mathematiktest erreichten Punkte

Aufgrund der Ergebnisse des Kolmogorov-Smirnov-Test ($p = 0.00$) wird eine Abweichung von der Normalverteilung angenommen, jedoch wird auf Grund der großen Stichprobenzahl auf den Einsatz von parameterfreien Verfahren verzichtet. Abbildung 7 zeigt die Gesamtverteilung der Punkte, welche von den Teilnehmern im Baseline Mathematiktest erreicht wurden. Durchschnittlich erreichten die

Teilnehmer eine Punktzahl von 3.05 mit einer Standardabweichung von 1.396 und einen Median von 3.

Tabelle 5 gibt die Aufgabenschwierigkeiten sowie den Prozentanteil der Personen, welche die Aufgaben richtig beantwortet haben, an. Die Aufgabenschwierigkeiten wurden mit einem Zwei Punkte Schema berechnet. Auf Grund dessen wurden die Teilpunkte, welche die Versuchspersonen erzielten, mit einbezogen. Tabelle 6 zeigt eine Reliabilitätstabelle des Pre-Priming Mathematiktest. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben. Der Reliabilitätswert würde sich nur bei der Elimination von Aufgabe 4 oder 7 geringfügig erhöhen.

Tabelle 5: Aufgabenschwierigkeiten der acht Aufgaben des Pre-Priming Mathematiktest

Aufgabe	1	2	3	4	5	6	7	8
% richtig gelöst	85.9	72.6	18.7	53.5	5	20.7	6.6	41.9
Aufgaben-Schwierigkeit	86	73	21.5	53.5	8.5	23.5	7.5	44

Aufgabe 1 wurde von den meisten Personen richtig beantwortet, während Aufgabe 5 und 7 am seltensten richtig beantwortet wurden.

Tabelle 6: Reliabilitätstabelle des Pre-Priming Mathematiktest, welcher nach dem Priming vorgegeben wurde. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben

Aufgabe	Skalenmittel, wenn Aufgabe gel.	Skalenvarianz, wenn Aufgabe gel.	Aufgaben-gesamtkorrelation	Cronbach's Alpha wenn Aufgabe gel.
1	2.191	1.688	0.152	0.392
2	2.324	1.536	0.191	0.374
3	2.863	1.710	0.083	0.423
4	2.515	1.401	0.251	0.340
5	3.000	1.767	0.230	0.377
6	2.842	1.542	0.239	0.351
7	2.983	1.900	-0.021	0.442
8	2.631	1.384	0.274	0.325
Rel.koeffizient	Alpha		0.413	

2.5.3 Mathematiktest – Post-Priming

Der verwendete Mathematiktest wurde nach dem Priming gestellt. Die maximale zu erreichende Punktezahl betrug bei diesem Test acht Punkte. Insgesamt bearbeiteten 225 Personen den vollständigen Mathematiktest (welcher nach dem Priming vorgegeben wurde) und können zur Berechnung herangezogen werden. Von den 225 Personen waren 119 weiblich und 106 männlich. Anhand des

Unterschieds, hinsichtlich der von den Teilnehmern erreichten Punktzahl bei dem Baseline Mathematiktest und bei diesem Mathematiktest, sollte der Effekt der unterschiedlichen Priming-Methoden gemessen werden. Betrachtet man die Geschlechter einzeln, erreichen die weiblichen Teilnehmer einen mittleren Punktwert von $M = 2.74$ (Median 3) und männliche Teilnehmer einen von $M = 2.75$ (Median 3) bei dem Mathematiktest, welcher nach dem Priming vorgegeben wurde. Die Deskriptiven Statistiken des Mathematiktest nach dem Priming finden sich in Tabelle 7. Der Test erreicht ein Cronbach's Alpha von $\alpha = 0.567$ (Tabelle 8).

Tabelle 7: Deskriptive Statistiken des Mathematiktest nach dem Priming

	Weibliche TP	Männliche TP	Gesamt
N	119	106	225
Mittelwert	2.744	2.745	2.742
Standard Abweichung	1.602	1.651	1.622
Minimale Punktzahl	0	0	0
Maximale Punktzahl	8	7	8

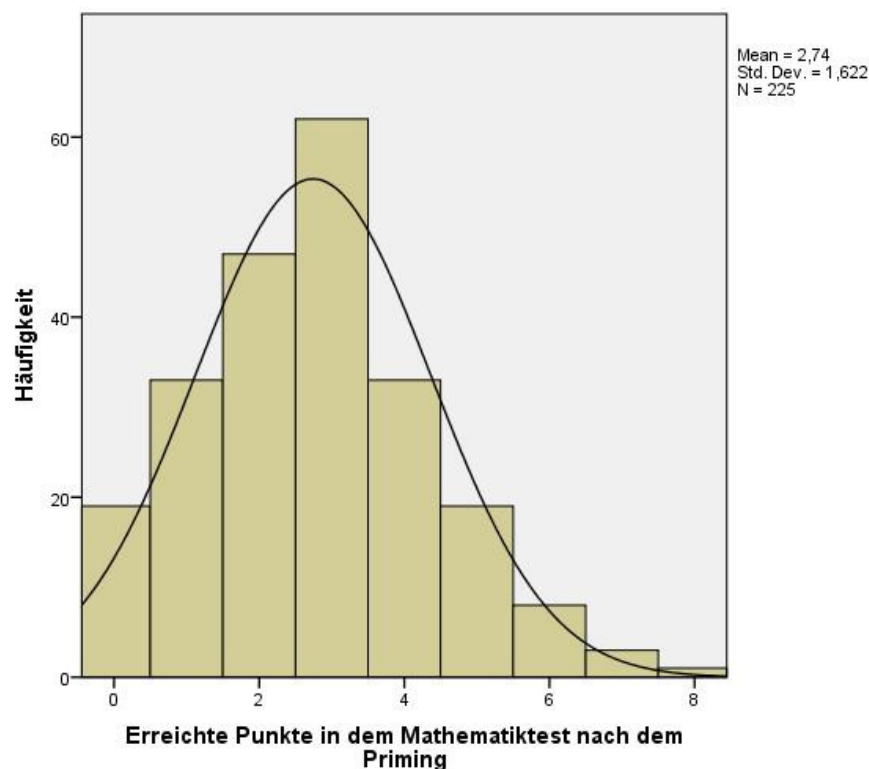


Abbildung 8: Ein Histogramm der Verteilung der von den Teilnehmern erreichten Punkte

Das Ergebnis des Kolmogorov Smirnov Tests ist signifikant ($p = 0.00$), daher wird eine Abweichung von der Normalverteilung angenommen. Abbildung 8 stellt die Verteilung der von den Schülern erreichten Punkte im Mathematiktest, welcher nach dem Priming vorgegeben wurde, graphisch dar. Dennoch wird im Folgenden auf Grund der hohen Stichprobengröße, auf den Einsatz parameterfreier Verfahren verzichtet.

Wie auch beim Baseline Mathematiktest wurden zwei Punkte für richtige Antworten und immer noch ein Punkt für den richtigen Rechenweg vergeben, jedoch wurde diese Wertung nur für die Aufgabenschwierigkeit (Tabelle 9) verwendet. Für alle anderen Berechnungen wurde eine Antwort entweder als richtig (ein Punkt), oder falsch (null Punkte) bewertet.

Tabelle 8: Reliabilitätstabelle des Post-Priming Mathematiktest. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben

Aufgabe	Skalenmittel, wenn Aufgabe gel.	Skalenvarianz, wenn Aufgabe gel.	Aufgaben-Gesamtkorrelation	Aufgaben-Gesamtkor. wenn Aufgabe gel.
1	2.022	2.134	0.221	0.553
2	2.147	1.957	0.314	0.521
3	2.364	1.983	0.288	0.546
4	2.191	2.075	0.205	0.565
5	2.671	2.338	0.246	0.546
6	2.524	2.001	0.370	0.501
7	2.702	2.380	0.285	0.543
8	2.640	2.160	0.386	0.508
Rel.koeffizient	Alpha		0.567	

Tabelle 9: Aufgabenschwierigkeit der acht Aufgaben des Post-Priming Mathematiktest

Aufgabe	1	2	3	4	5	6	7	8
% Richtig gelöst	73.3	60.4	38.7	56.0	8.0	22.7	4.9	11.1
Aufgabenschwierigkeit	73.5	61	40.5	56	8.5	24	5	11

2.5.4 Kurzfragebogen zu den Arbeitshaltungen

Zur Analyse des kurzen Fragebogens, über den eigenen Arbeitsstil bei der Bearbeitung der Aufgaben, wurden die semantischen Differenziale zunächst einheitlich ausgerichtet, sodass ein hoher Wert für eine positiv besetzte Antwort steht. Die in der Folge verwendete Aufgabenbezeichnung bezieht sich auf diesen positiven Pol. Eine anschließende Faktorenanalyse mit Varimax- Rotation liefert zwei Faktoren (Anhang E, Tabelle 40).

- Auf Faktor 1 (erklärt 37 % der Varianz) laden die Adjektive „Bemühung“, „Konzentration“, „Interesse“ und „Motivation“. Dieser Faktor kann als Maß für die generelle Arbeitsmotivation zum Testzeitpunkt gesehen werden.
- Auf Faktor 2 (erklärt 25 % der Varianz) laden „Schwierigkeit“, „Forderung“ und „Geschwindigkeit“. Dieser Faktor kann als generelles Maß für die empfundene Schwierigkeit der Aufgaben gesehen werden.

Bei dem zweiten Testzeitpunkt zeigt sich ein ähnliches Bild. Die Faktorenanalyse ergibt ebenfalls zwei Faktoren.

- Auf Faktor 1 (erklärt 39 % der Varianz) laden die Adjektive „Bemühung“, „Konzentration“, „Interesse“ und „Motivation“. Dieser Faktor kann als Maß für die generelle Arbeitsmotivation zum Testzeitpunkt zwei gesehen werden.
- Auf Faktor 2 (erklärt 25 % der Varianz) laden „Schwierigkeit“, „Forderung“ und „Geschwindigkeit“ zum Testzeitpunkt zwei. Dieser Faktor kann als generelles Maß für die empfundene Schwierigkeit der Aufgaben gesehen werden.

3 Ergebnisse

Im folgenden Kapitel kommen bestimmte Abkürzungen statistischer Kennwerte zum Einsatz. Dabei steht N für den Stichprobenumfang, M für den Mittelwert, SD für die Standard Abweichung, p für die Signifikanz, df für die Freiheitsgrade, MQ für das Mittel der Quadrate, F für die Prüfgröße der F-Verteilung und r_s für den Spearman'schen Korrelationskoeffizient.

3.1 Haupthypothesen

Obwohl eine Abweichung von der Normalverteilung in der Stichprobe unserer Teilnehmer, bezüglich der erreichten Punkte in beiden Mathematiktests vorhanden ist, wurde wenn möglich auf den Einsatz parameterfreier Verfahren verzichtet. Da unsere Stichprobe einen ausreichend großen Umfang besitzt, kamen zur Hypothesenprüfung zumeist t-Tests und Varianzanalysen zum Einsatz.

3.1.1 H1: Gibt es einen signifikanten Unterschied in der Basisfertigkeit Mathematik zwischen Mädchen und Burschen?

Um die Basisfertigkeit für Mathematik zu bestimmen wird eine Summe aus den Punkten, die die Teilnehmer beim ersten Termin (Mathematik Test Baseline Teil 1) und beim zweiten Termin (Mathematiktest Baseline Teil 2) erreicht haben, gebildet.

Um zu überprüfen, ob Männer und Frauen sich hinsichtlich ihrer grundlegenden Fähigkeiten in Mathematik signifikant unterscheiden, wird ein t-Test für unabhängige Stichproben durchgeführt (Tabelle 10). Die Basisfähigkeit Mathematik wird hier durch die erreichte Punktezahl im Mathematiktest, welcher vor dem Priming bearbeitet wurde, repräsentiert. Das Ergebnis des Levene's Test zeigt, dass die Homogenität der Varianzen nicht angenommen werden kann, dies wird bei der Interpretation des t-Test Ergebnisses berücksichtigt. Der t-Test ergibt ein $t(239) = -0.409$ ($p = 0.072$). Es wird daher angenommen, dass es keinen signifikanten Unterschied in den erreichten Punkten im Mathematiktest vor dem Priming zwischen Mädchen und Jungen gibt.

Tabelle 10: Deskriptive Statistiken und Ergebnisse des t-Test für unabhängige Stichproben und Levene's Test der weiblichen und männlichen Teilnehmer bezüglich des Pre-Priming Mathematiktest

t-Test für unabhängige Stichproben							
Geschlecht	N	Mittelwert	Std. Abweichung	Std. Fehler Mittelwert	t	df	Sig.
Weiblich	129	2.899	1.286	0.113	-1.806	239	0.072
Männlich	112	3.223	1.499	0.141			
Levene's Test				F	df		Sig.
				7.996	220,209		0.005

Im Vergleich der Mathematiknoten (Tabelle 11) zeigt sich, dass ein hoher Prozentsatz beider Geschlechter angeben, gute Note (1 – 2) in Mathematik zu haben. Relativ wenige Personen geben an, eine 4 oder 5 in Mathematik bekommen zu haben und sogar mehr männliche als weibliche Teilnehmer gaben an, eine 4 in Mathematik gehabt zu haben. Der Median der angegebenen Schulnoten in Mathematik liegt bei beiden Geschlechtern bei 3. Ein U-Test $z = -1.237$ ($U = 6581$, $p = 0.216$) zeigt keinen signifikanten Unterschied zwischen den Geschlechtern bezüglich der angegebenen Mathematiknote. Tabelle 12 zeigt die Rangdaten der Mathematiknoten.

Tabelle 11: Angegebene Mathematiknoten der männlichen und weiblichen Teilnehmer im Vergleich

Note	Weibliche Teilnehmer	Männliche Teilnehmer
	Prozent	Prozent
1	26.9	23.6
2	19.3	17.9
3	37.8	34.9
4	10.9	21.7
5	5	1.9

Tabelle 12: Rangdaten der Mathematiknoten

Geschlecht	N	Mittlerer Rang	Rangsumme	Median
Weiblich	129	116.016	14966	3
Männlich	112	126.741	14195	3
Gesamt	241			3

Es wurde eine Spearman Korrelation gerechnet, um den Zusammenhang zwischen den angegebenen Noten und den Testergebnissen im Mathematiktest zu untersuchen. Es ergibt sich ein Wert von $r_s = -0.211$, welcher einen

Signifikanzwert von $p = 0.01$ aufweist. Daher kann man davon ausgehen, dass ein geringer negativer Zusammenhang zwischen den angegebenen Noten und den Ergebnissen im Baseline Mathematiktest vorhanden ist. Nicht überraschend stellte sich heraus, dass Personen die, angaben eine gute Note in Mathematik zu haben, auch tendenziell eher bessere Leistungen im Baseline Mathematiktest erbrachten.

3.1.2 CFT – Unterschiede in der Fähigkeit des logischen Denkens

Teilnehmer.

Unterschiede zwischen den Geschlechtern bezüglich der mittels CFT gemessenen Fähigkeit im logischen Denken der Teilnehmer wurden nicht angenommen, jedoch soll durch einen Vergleich sichergestellt werden, dass keine solchen Unterschiede bestehen. Es wird ein t-Test für unabhängige Stichproben durchgeführt um zu überprüfen, ob sich Unterschiede in der Fähigkeit zum logischen Denken zwischen den Geschlechtern finden. Die Ergebnisse des t-Tests finden sich in Tabelle 13. Es besteht laut t-Test kein signifikanter Unterschied zwischen den Geschlechtern, bezüglich der gemessenen Grundintelligenz.

Tabelle 13: Ergebnisse des t-Test und des Levene's Test für die erreichten Punkte im CFT Untertest Matrizen

t-Test für unabhängige Stichproben							
Geschlecht	N	Mittelwert	Std. Abweichung	Std. Error Mittelwert	T	df	Sig.
Weiblich	129	5.667	1.876	0.165	-0.042	239	0.966
Männlich	112	5.679	2.398	0.226			
Levene's Test				F	df		Sig.
				10.993	209.118		0.001

3.1.3 Haupteffekt des Primings.

Bei der Untersuchung der Hypothese wurden 225 Versuchspersonen miteinbezogen, da diese einen kompletten Baseline Mathematiktest (vor dem Priming) sowie einen kompletten Mathematiktest, welcher nach dem Priming vorgegeben wurde, absolviert hatten. Aufgrund der großen Anzahl der Teilnehmer wird kein parameterfreies Verfahren, sondern eine Varianzanalyse mit Messwiederholung durchgeführt, um den Effekt des Primings zu überprüfen. Tabelle 14 zeigt die Ergebnisse der Varianzanalyse mit Messwiederholung. Die

Varianzanalyse ergibt, dass es einen signifikanten Unterschied zwischen der Leistung im Pre-Priming-Test und der Leistung im Mathematiktest, welcher nach dem Priming vorgegeben wurde, gibt. Es kommt zu einer signifikanten Minderung der Leistung im Vergleich von dem Baseline zu dem Mathematiktest. Weder das Geschlecht noch das Priming hat einen signifikanten Einfluss auf die Leistung. Auch die Wechselwirkungen sind nicht signifikant. Der Begriff „Test“ in der Tabelle 14 meint in diesem Fall, die erreichte Punktezahl im Baseline Mathematiktest (vor dem Priming) bzw. der erreichten Punktezahl nach dem Priming.

Tabelle 14: Varianzanalyse mit Messwiederholung und Levene's Test der Mathematiktests unter Berücksichtigung des Primings und des Geschlechts aller Teilnehmer mit vollständigen Datensets

Varianzanalyse mit Messwiederholung				
Effekt	F	Sig.	MQ	df
Test	13.021	0.000	14.362	1
Geschlecht	1.120	0.291	3.944	1
Priming	0.003	0.953	0.012	1
Geschlecht x Priming	0.921	0.338	0.245	1
Test x Priming	1.157	0.283	1.276	1
Test x Geschlecht	3.163	0.077	3.489	1
Test x Priming x Geschlecht	0.054	0.816	0.060	1
Levene Test	F	Sig	df1	df2
Pre-Priming Mathematiktest	3.137	0.026	3	221
Post – Priming Mathematiktest	3.086	0.028	3	221

Levene's Test weist keine Homogenität der Varianzen für den Baseline Mathematiktest (Pre-Priming Mathetest), das gleiche gilt für den nach dem Priming bearbeiteten Mathematiktest (Post-Priming Mathematiktest) auf. Da aber bereits bei der Datenerhebung auf eine ausgewogene Verteilung der Personen auf die einzelnen Versuchsbedingungen geachtet wurde und die Stichprobe einen ausreichend großen Umfang aufweist, wird die Varianzanalyse mit Messwiederholung dennoch gerechnet. Die Mittelwerte sowie der Umfang der einzelnen Versuchsgruppen sind in Tabelle 15 zu finden.

Tabelle 15: Mittelwerte und Standardabweichungen der erreichten Punkte im Mathematiktest vor und nach dem Priming aufgeteilt auf ihre Primingbedingungen

	Geschlecht	Priming	Mittelwert	St. Abweichung	N
Erreichte Punkte Pre-Priming Mathematiktest	Weiblich	Stereotypnonkonform	3.066	1,263	61
		Stereotypkonform	2.776	1,325	58
		Gesamt	2.924	1,296	119
	Männliche	Stereotypnonkonform	3.236	1.374	55
		Stereotypkonform	3.333	1.657	51
		Gesamt	3.283	1.510	106
	Gesamt	Stereotypnonkonform	3.147	1.314	116
		Stereotypkonform	3.037	1.509	109
		Gesamt	3.093	1.410	225
Erreichte Punkte Post-Priming Mathematiktest	Weiblich	Stereotypnonkonform	2.754	1.468	61
		Stereotypkonform	2.724	1.745	58
		Gesamt	2.739	1.602	119
	Männlich	Stereotypnonkonform	2.618	1.881	55
		Stereotypkonform	2.882	1.366	51
		Gesamt	2.745	1.651	106
	Gesamt	Stereotypnonkonform	2.690	1.670	116
		Stereotypkonform	2.798	1.574	109
		Gesamt	2.742	1.622	225

Wenn man die Teilnehmer aus den Berechnungen ausschließt, welche die Frage zum Inhalt des Artikels falsch beantwortet haben (und ihn daher höchstwahrscheinlich nicht gelesen haben), zeigt sich ein signifikanter Effekt bei dem Geschlecht. Tabelle 16 zeigt das Ergebnis der Varianzanalyse mit Messwiederholung für die Teilnehmer, die die Frage zum Inhalt des Artikels richtig beantwortet haben. Der Levene's Test (Tabelle 16) ergibt hier für den Mathematiktest vor dem Priming, sowie für den Mathematiktest nach dem Priming kein signifikantes Ergebnis, daher kann die Homogenität der Varianzen angenommen werden.

Tabelle 16: Varianzanalyse mit Messwiederholung und Levene's Test unter Ausschluss der Personen, welche die Frage nach dem Inhalt des Artikels falsch beantworteten

Varianzanalyse mit Messwiederholung				
Effekt	F	Sig.	MQ	df
Test	6.814	0.010	7.327	1
Geschlecht	6.067	0.015	20.162	1
Priming	0.830	0.773	0.276	1
Geschlecht x Priming	1.126	0.290	3.741	1
Test x Priming	0.084	0.772	0.090	1
Test x Geschlecht	0.306	0.594	0.306	1
Test x Priming x Geschlecht	0.336	0.563	0.362	1
Levene Test	F	Sig	df1	df2
Pre-Priming Mathematiktest	2.388	0.070	3	220
Post – Priming Mathematiktest	2.551	0.057	3	220

Betrachtet man die deskriptiven Statistiken, erkennt man einen Unterschied in den Mittelwerten von weiblichen und männlichen Teilnehmern (Tabelle 17). Ein t-Test für abhängige Stichproben wurde durchgeführt, um die Geschlechter einzeln zu betrachten. Der t-Test für abhängige Stichproben ergibt ein signifikantes Ergebnis für die männlichen Teilnehmer mit einem $t(88) = 2.269$ ($p = 0.026$), jedoch kein signifikantes Ergebnis für die weiblichen Teilnehmer mit einem $t(114) = 1.516$ ($p = 0.132$). Daraus folgt, dass sich die Leistung der Männer von dem ersten zum zweiten Test signifikant verändert hat.

Tabelle 17: Deskriptive Statistiken der erreichten Punkte im Pre-Priming Mathematiktest und Post-Priming Mathematiktest unter Ausschluss der Personen, welche die Frage nach dem Inhalt des Artikels falsch beantworteten

Geschlecht	Test	N	Min	Max	Mittelwert	St. Abweichung
Weiblich	Pre-Priming Mathematiktest	115	0	6	2.858	1.304
	Post-Priming Mathematiktest		0	8	2.687	1.564
Männlich	Pre-Priming Mathematiktest	89	0	7	3.404	1.520
	Post- Priming Mathematiktest		0	7	3.078	1.539

3.1.4 H2: Weibliche Versuchspersonen, welche den stereotypnonkonformen Artikel (SL) bekommen, erbringen bessere oder gleich gute Leistungen bei dem Mathematiktest nach dem Priming wie im Mathematiktest vor dem Priming.

Unsere erste Hypothese war, dass weibliche Teilnehmer, welche mittels des Artikels geprimt wurden, jener die Information enthielt, dass Mädchen besser in mathematischen Bereichen sein sollten, dem Stereotyp Lift unterliegen und demnach bessere oder zumindest gleiche Leistungen wie in dem Baseline Test erbringen.

Tabelle 18: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	N	Mittelwert	St. Abweichung	Min.	Max
Pre-Priming Mathematiktest	61	3.066	1.263	1	6
Post-Priming Mathematiktest		2.754	1.468	0	6

Wie in Tabelle 18 zu sehen ist, haben die weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen im Baseline einen höheren Mittelwert als im Mathematiktest, welcher nach dem Priming vorgegeben wurde. Ein t-Test für abhängige Stichproben zeigt mit einem $t(60) = 1.575$ ($p = 0.120$), dass der Leistungsabfall von dem Baseline Mathematiktest zum Mathematiktest, welcher nach dem Priming vorgegeben wurde, nicht signifikant ausfällt. Daher verschlechtern sich die weiblichen Teilnehmer dieser Bedingung von dem Baseline zu dem Mathematiktest (welcher nach dem Priming vorgegeben wurde), jedoch nicht signifikant. Wir kommen zu einem ähnlichen Ergebnis wie Johnson et al. (2012), mit dem Unterschied das sich unsere Stichprobe verschlechtert. Die Hypothese kann daher zumindest Teilweise als bestätigt angesehen werden. 97 % der Teilnehmer dieser Teilstichprobe hatten die Frage nach dem Inhalt des Artikels richtig beantwortet und 62.3 % glaubten, dass die Kernaussage des Artikels stimmt. Die Teilnehmer wurden gefragt, für wie glaubwürdig sie den Artikel hielten. Tabelle 19 zeigt die Verteilung. Trotz einer hohen Zustimmung und Einschätzung der Glaubwürdigkeit des Artikelinhalts, hat dieser keinen Effekt auf die Leistungen dieser Versuchspersonen.

Tabelle 19: Glaubwürdigkeitseinschätzung des Artikels seitens der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	Prozent	Häufigkeit
Geringe Glaubwürdigkeit (0-2)	8.2	5
Mittlere Glaubwürdigkeit (3-4)	65.6	40
Hohe Glaubwürdigkeit (5-6)	26.2	16

Die Motivation der Teilnehmer bei der Bearbeitung des Mathematiktest (nach dem Priming), unabhängig von dem Zeitpunkt der Testung, war größtenteils durchschnittlich (Median = 3) mit einer Tendenz zur Mitte.

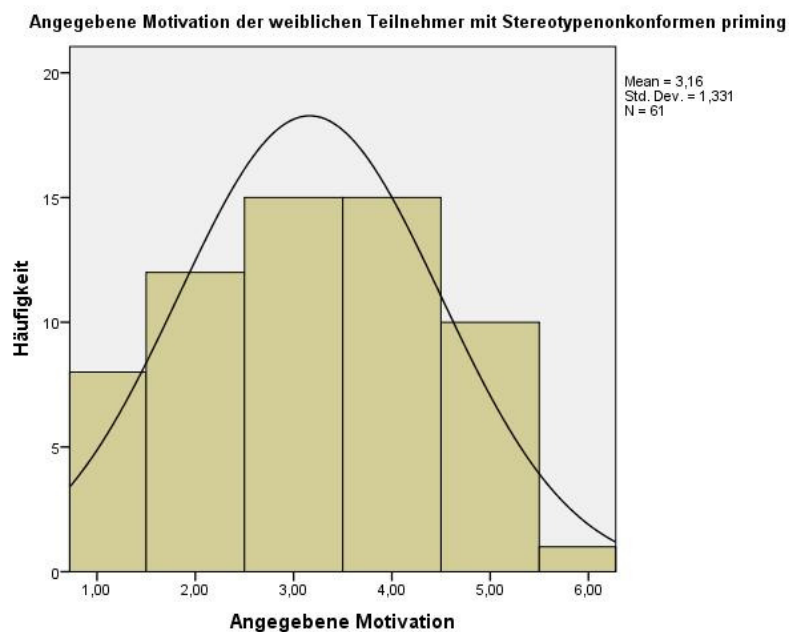


Abbildung 9: Histogramm, welches die Motivation der weiblichen Teilnehmer mit stereotypnonkonformen Priming zeigt

Die Teilnehmer konnte ihre Motivation auf einer Skala von 1 bis 6 einschätzen. Acht Personen gaben an, eine Motivation von 1 zu haben. Da die Fragen bezüglich der Arbeitshaltung ganz zum Schluss eines jeden Testtermins erfolgten, bezieht sich die angegebene Motivation auf den gesamten Termin. Die angegebene Motivation (Abbildung 9) bezieht sich außerdem auf die Bearbeitung aller Aufgaben, welche Personen bei dem Termin bei dem der Mathematiktest, welcher das Priming mittels Artikel nutze, bearbeitet hatten.

Bezüglich der Angaben zu den anderen Arbeitshaltungen ergab sich das in Tabelle 20 angeführte Bild.

Tabelle 20: Mediane und Quartile der Arbeitshaltungen der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	Median	Quartile	
		25	75
Konzentration	3	2	4
Forderung	3	3	4
Geschwindigkeit	4	3	4
Interesse	4	2	4
Bemühung	4	3	4
Schwierigkeit	3	3.5	4

3.1.5 H3: Weibliche Versuchsteilnehmer welche einen stereotypkonformen Artikel (Stereotype Threat) erhalten, bringen im Mathematiktest nach dem Priming schlechtere Leistungen als im Mathematiktest vor dem Priming.

Wir erwarteten eine Verschlechterung der Leistungen jener weiblicher Probanden, welche den Artikel mit der Information, dass Männer besser in mathematischen Bereichen seien, erhielten. Ein t-Test für abhängige Stichproben ergibt ein $t(57) = 0.256$ ($p = 0.799$) und zeigt kein signifikantes Ergebnis. Tabelle 21 zeigt die deskriptiven Statistiken der Mathematiktests vor und nach dem Priming für die weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefen. Die weiblichen Teilnehmer, welche den stereotypkonformen Artikel erhielten, brachten im Post-Priming Mathematiktest nicht signifikant schlechtere Leistungen als in dem Baseline Mathematiktest.

Tabelle 21: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	N	Mittelwert	St. Abweichung	Min.	Max
Pre-Priming Mathematiktest	58	2.775	1.325	0	6
Post-Priming Mathematiktest	58	2.724	1.745	0	8

Nur zwei Personen beantworteten die Frage nach der Kernaussage des Artikels falsch und 67.2 % der Teilnehmer dieser Priminggruppe glaubten, dass die Inhalte des Artikels stimmen. Tabelle 22 zeigt, dass ein höherer Prozentsatz der

Teilnehmer die Glaubwürdigkeit des stereotypkonformen Artikel als gering einschätzt als bei dem stereotypnonkonformen.

Tabelle 22: Glaubwürdigkeitseinschätzung des Artikels seitens der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	Prozent	Häufigkeit
Geringe Glaubwürdigkeit (0-2)	20.7	12
Mittlere Glaubwürdigkeit (3-4)	53.4	31
Hohe Glaubwürdigkeit (5-6)	25.9	15

In dieser Versuchsgruppe beschrieben die Teilnehmer ihre Motivation auch größtenteils als durchschnittlich, mit einem Median von 3, wobei neun Personen ihre Motivation mit 1 angaben. Abbildung 10 stellt die Verteilung der Motivationseinschätzungen graphisch dar. Eine Person füllte den Fragebogen zu den Arbeitshaltungen nicht vollständig aus und wurde daher ausgeschlossen.

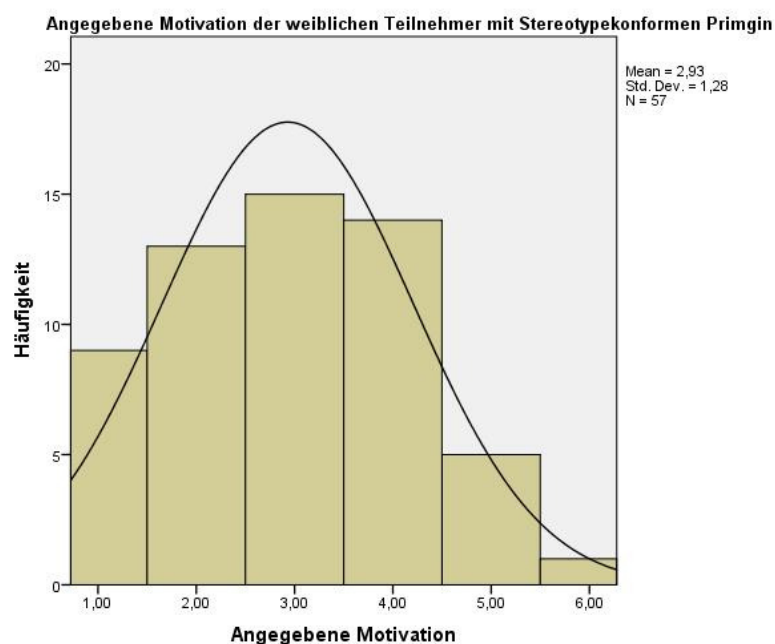


Abbildung 10: Histogramm, welches die Motivation der weiblichen Teilnehmer mit stereotypkonformem Priming zeigt

Die Ergebnisse der anderen Arbeitshaltungen für diese Gruppe werden in Tabelle 23 dargestellt.

Tabelle 23: Mediane und Quartile der Arbeitshaltungen der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	Median	Quartile	
		25	75
Konzentration	3	2	4
Forderung	3	2	4
Geschwindigkeit	4	3	5
Interesse	3	3	4
Bemühung	3	2	4
Schwierigkeit	3	2	5

3.1.6 H4: Weibliche Teilnehmer schneiden im Mathematiktest nach dem Priming unter Stereotype Lift besser ab, als unter Stereotype Threat.

Der Unterschied in den Leistungen zwischen den beiden Bedingungen ist so gering, dass man diese Hypothese als verworfen ansehen kann. Weibliche Teilnehmer schnitten in beiden Bedingungen beinahe gleich gut ab. Mit einem Mittelwert von 2.75 ($SD = 1.468$) bei den weiblichen Teilnehmern mit stereotypenonkonformen Priming und einem Mittelwert von 2.72 ($SD = 1.745$) bei den weiblichen Teilnehmern mit stereotypkonformen Priming, ergibt sich nur ein Mittelwertsunterschied von 0.03. Ein t-test für unabhängige Stichproben (Homogenität der Varianzen kann angenommen werden, siehe Anhang E, Tabelle 45) mit $t(117) = 0.102$ ($p = 0.917$) bestätigt diese Annahme.

3.1.7 H5: Männliche Versuchsteilnehmer, welche der Stereotype Lift Bedingung ausgesetzt werden, bringen genauso gute Leistungen im Mathematiktest vor, wie im Mathematiktest nach dem Priming.

Von männlichen Versuchsteilnehmern, welche den stereotypkonformen Artikel gelesen hatten, wurde erwartet, dass sich ihre Leistung nicht verändern würde. Ein t-Test für abhängige Stichproben zeigt mit einem $t(50) = 2.602$ ($p = 0.012$), dass auch diese Teilgruppe der Versuchsteilnehmer signifikant schlechtere Leistungen beim Mathematiktest, welcher nach dem Priming vorgegeben wurde,

erbringt. Tabelle 24 gibt die deskriptiven Statistiken, sowie das Ergebnis des t-Tests der männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefen wieder.

Tabelle 24: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming, sowie das Ergebnis des t-Test der männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	N	Mittelwert	St. Abweichung	Min.	Max
Pre-Priming Mathematiktest	51	3.333	1.657	0	7
Post-Priming Mathematiktest	51	2.882	1.366	0	6
t-Test	t	2.602	Sig. 0.012		

Von den 51 Teilnehmern dieser Teilstichprobe räumten 6 (11.8 %) dem Artikel eine geringe Glaubwürdigkeit ein. Tabelle 25 zeigt die durch die Teilnehmer empfundene Glaubwürdigkeit des Artikels. Eine Person dieser Teilstichprobe hat diese Frage nicht beantwortet, möglicherweise brach diese Person die Befragung früher ab.

Tabelle 25: Glaubwürdigkeitseinschätzung des Artikels seitens der männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	Prozent	Häufigkeit
Geringe Glaubwürdigkeit (0-2)	11.7	6
Mittlere Glaubwürdigkeit (3-4)	62.8	32
Hohe Glaubwürdigkeit (5-6)	23.5	12

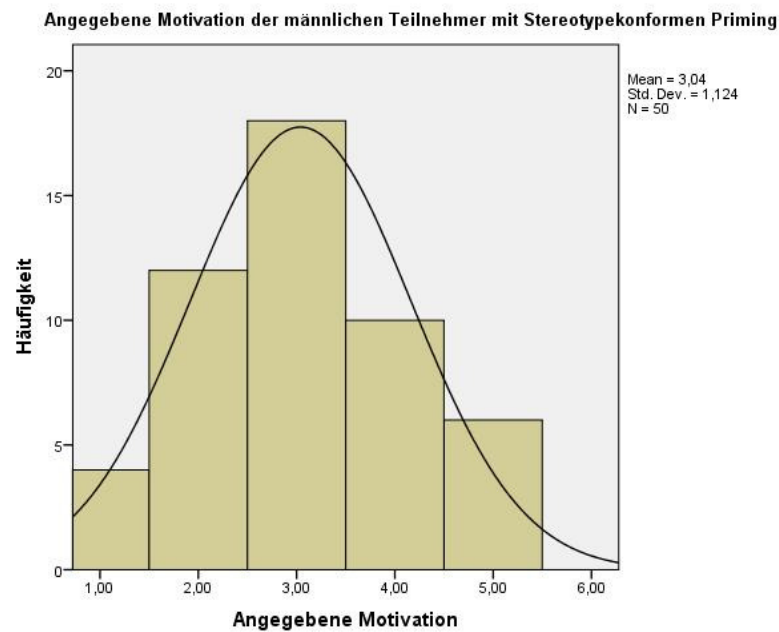


Abbildung 11: Histogramm, welches die Motivation der männlichen Teilnehmer mit stereotypkonformen Priming zeigt

Abbildung 11 zeigt ein Histogramm der eingeschätzten Motivation seitens der männlichen Teilnehmer, welche das stereotypkonforme Priming erhielten. Die eingeschätzte Motivation ist jene zu dem Testzeitpunkt, an welchem sie das Priming mittels Artikel erhielten. Nur vier Personen gaben in dieser Gruppe an, eine Motivation von 1 zu haben, damit ist dies die Gruppe mit den wenigsten (*gar nicht*) motivierten Personen. Diese Gruppe hat bei der Motivation einen Median von 3.

Die Ergebnisse der anderen Arbeitshaltungen finden sich in Tabelle 26.

Tabelle 26: Mediane und Quartile der Arbeitshaltungen männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefen

	Median	Quartile	
		25	75
Konzentration	3	2	4
Forderung	3	2	4
Geschwindigkeit	4	3	5
Interesse	3	3	4
Bemühung	3	2	4
Schwierigkeit	3	2	5

3.1.8 H6: Männliche Versuchsteilnehmer welche der Stereotype Threat Bedingung ausgesetzt werden, bringen im Mathematiktest nach dem Priming bessere Leistungen als im Mathematiktest vor dem Priming.

Es wurde erwartet das männliche Teilnehmer, welche den Artikel mit der Kernaussage bekamen „Frauen erbringen bessere Leistungen in mathematischen Bereichen“ bessere Leistungen im Mathematiktest nach dem Priming erbringen würden, als im Baseline Test. H. Johnson et al. (2012) hatten bei Ihrer Untersuchung einen solchen Effekt beobachten können. Wir können diesen Effekt in dieser Untersuchung nicht nachstellen. Ein t-Test für abhängige Stichproben zeigt mit einem $t(54) = 2.921$ ($p = 0.005$), dass sich die Teilnehmer nicht signifikant verbessert, sondern signifikant verschlechtert haben. 20 % der Teilnehmer beantworteten die Frage nach der Hauptaussage des Artikels falsch. Tabelle 27 gibt an, für wie Glaubwürdig die Teilnehmer die Inhalte des Artikels halten. Auch in dieser Gruppe schreiben die meisten Personen dem Artikel eine mittlere Glaubwürdigkeit zu.

Tabelle 27: Deskriptive Glaubwürdigkeitseinschätzung des Artikels seitens der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	Prozent	Häufigkeit
Geringe Glaubwürdigkeit (0-2)	23.7	13
Mittlere Glaubwürdigkeit (3-4)	63.7	35
Hohe Glaubwürdigkeit (5-6)	12.7	7

Beim Baseline Test hatten die Teilnehmer einen Punktedurchschnitt von 3.24 mit einer Standardabweichung von 1.37 und beim Mathematiktest nach dem Priming nur noch einen Punktedurchschnitt von $M = 2.62$ mit einer Standardabweichung von 1.88. Tabelle 28 enthält die Ergebnisse des t-Test sowie die deskriptiven Statistiken dieser Teilstichprobe.

Tabelle 28: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming sowie das Ergebnis des t-Test der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	N	Mittelwert	St. Abweichung	Min.	Max
Pre-Priming Mathematiktest	55	3.236	1.374	1	5
Post-Priming Mathematiktest	55	2.618	1.881	0	7
t-Test	t	2.921	Sig. 0.005		

Die Motivation der Teilnehmer in dieser Versuchsgruppe war ebenfalls größtenteils durchschnittlich, mit einem Median von 3 (Quartil 25 =, Quartil 75 = 4). Abbildung 12 zeigt ein Histogramm der Einschätzung der Teilnehmer bezüglich ihrer Motivation zu dem Termin, an dem sie den Mathematiktest mit dem Priming durch den Artikel bearbeiteten. Mit 10 Personen ist dies die Gruppe mit den meisten Personen, welche ihre Motivation mit einer 1 einschätzten.

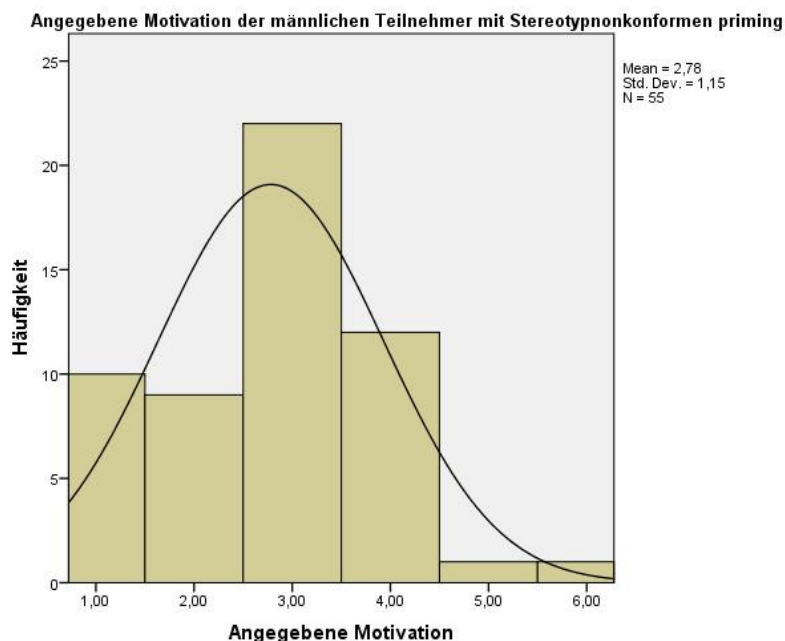


Abbildung 12: Histogramm welches die Motivation der männlichen Teilnehmer mit stereotypnonkonformen Priming zeigt

Bezüglich der anderen Arbeitshaltungen, ergibt sich folgendes Bild (siehe Tabelle 29).

Tabelle 29: Median und Quartile der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefen

	Median	Quartile	
		25	75
Konzentration	3	2	4
Forderung	3	3	3
Geschwindigkeit	4	3	5
Interesse	4	3	4
Bemühung	3	3	4
Schwierigkeit	4	3	4

3.1.9 H7: Männliche Versuchsteilnehmer in der Stereotype Threat Bedingung schneiden im Mathematiktest nach dem Priming, gleich gut ab wie männliche Versuchsteilnehmer in der Stereotype Lift Bedingung.

Männliche Probanden, sowohl in der ST, wie auch in der SL Bedingung, verschlechterten ihre Leistungen im Vergleich vom Baseline zum Mathematiktest, welcher nach dem Priming vorgegeben wurde, signifikant. In der ST Bedingung erreichten die Teilnehmer im Mathematiktest nach dem Priming einen Punktedurchschnitt von $M = 2.618$, mit einer $SD = 1.881$. In der SL Bedingung erreichten die Probanden einen Punktedurchschnitt von $M = 2.882$, mit einer Standardabweichung von 1.366. Ein T-Test für unabhängige Stichproben zeigt kein signifikantes Ergebnis $t(98.494) = -0.832$ ($p = 0.408$). Auch vor dem Priming (bei dem Baseline Mathematiktest) zeigte sich bei einem t-Test für unabhängige Stichproben (Homogenität der Varianzen konnte hier angenommen werden), kein signifikanter Unterschied ($t(104) = -0.329$, $p = 0.745$) zwischen den Leistungen der männlichen Probanden in den zwei verschiedenen Versuchsgruppen. Die Ergebnisse der entsprechenden Levene's Tests finden sich im Anhang E (Tabelle 44).

3.1.10 H8: Männliche Versuchsteilnehmer schneiden unter ST besser ab als Frauen unter ST.

Weibliche Versuchsteilnehmer in der ST Bedingung erreichten in dem Baseline Mathematiktest einen Mittelwert von 2.776 (*SD* von 1.325) und in dem nach dem Priming bearbeiteten Mathematiktest einen Mittelwert von 2.724 (*SD* von 1.745), während männliche Teilnehmer in der ST Bedingung Baseline Test einen Punktedurchschnitt von 3.236 mit einer *SD* = 1.347. Beim Mathematiktest, welcher nach dem Priming bearbeitet wurde erreichen die Teilnehmer einen Punktedurchschnitt von $M = 2.618$ mit einer Standardabweichung von 1.347 erreichten. Ein t-Test für unabhängige Stichproben (Homogenität der Varianzen wird angenommen) zeigt mit einem $t(111) = -0.311$ ($p = 0.757$) kein signifikantes Ergebnis zwischen diesen beiden Gruppen, in dem nach dem Priming bearbeiteten Mathematiktest. In Tabelle 30 sind die genauen Ergebnisse des t-Tests für unabhängige Stichproben abzulesen. Auch vor dem Priming zeigte sich kein Unterschied zwischen diesen Gruppen, bezüglich ihrer Mathematischen Leistungen im Baseline Mathematiktest.

Tabelle 30: Ergebnisse des t-Tests und Levene's Test von männlichen und weiblichen Teilnehmern, welche das ST Priming erhalten hatten

	Pre-Priming Mathematiktest	Post-Priming Mathematiktest
T	-1.812	0.310
Sig.	0.073	0.757
Df	111	111
Levene's Test	F	Sig
	1.324	0.252

Da der t-Test nicht signifikant ausfällt, kann diese Hypothese nicht als bestätigt angesehen werden. Zwar erreichten die männlichen Teilnehmer im Post-Priming Mathematiktest einen höheren durchschnittlichen Punktwert als die weiblichen Teilnehmer, jedoch ist dieser Unterschied nicht signifikant.

3.2 Weitere Fragestellungen

Da unsere Nebenhypothesen alle mit der Stärke und Richtung des Primingeffektes zusammenhängen und dieser nicht aufgetreten ist, werden sie verworfen. Jene, welche sich nicht mit dem Effekt des Priming auf die Ergebnisse der Mathematiktests befassen, werden im weiteren Verlauf besprochen.

3.2.1 Hat Interesse an Mathematik einen Einfluss auf Leistungen im Mathematiktest, welcher vor dem Priming bearbeitet wurde

Eine Spearman Korrelation zwischen dem Interesse an Mathematik und dem erreichten Punktwert in dem Baseline Mathematiktest zeigt ein signifikantes Ergebnis von $r_s = 0.22$. Dieser Zusammenhang ist auf einem Niveau von $p = 0.001$ signifikant, jedoch sehr gering.

3.2.2 Hat die persönliche Wichtigkeit von Mathematik einen Effekt auf die Leistungen im Mathematiktest welcher vor dem Priming bearbeitet wurde?

Eine Spearman Korrelation zwischen dem angegebenen Punktwert des Elements *persönliche Wichtigkeit von Mathematik* und dem erreichten Punktwert im Baseline Test wurde gerechnet. Das Ergebnis von $r_s = 0.19$ ist bei einem Signifikanzniveau von $p = 0.01$ höchst signifikant, jedoch ist der Zusammenhang sehr gering. Betrachtet man beide Geschlechter getrennt (Tabelle 31) wird die Korrelation nur bei den männlichen Teilnehmern signifikant.

Tabelle 31: Ergebnis der einseitigen Spearman Korrelation zwischen der persönlichen Wichtigkeit und der Einschätzung der erreichten Punkte im Baselinetest Mathematiktest

	Spearman Korrelation (r)
Weiblich	0.136
Männlich	0.245
Gesamt	0.19

3.2.3 Beeinflusst das Priming die Selbsteinschätzung?

Nach jedem Mathematiktest folgte eine Befragung, bei der Schüler einschätzen konnten, wie viele der Mathematikaufgaben sie richtig beantwortet hatten. Der Baseline Mathematiktest (vor dem Priming), war auf zwei Termine mit jeweils vier Mathematikaufgaben aufgeteilt. Nach jeder der beiden Hälften folgte eine Befragung nach den richtig beantworteten Mathematikaufgaben. Pro Termin war die Antwortmöglichkeit bei den Mathematiktesthälften des Pre-Priming Mathematiktest auf 0 bis 4 beschränkt. Die Einschätzungen beider Termine wurden addiert um die Selbsteinschätzungen für alle acht Aufgaben des Pre-Priming Mathematiktests zu erhalten. Bei dem Mathematiktest, welcher nach dem Priming zu bearbeiten war, konnten die Teilnehmer bei der Nachbefragung Zahlen von 0 bis 8 eintragen.

Um zu untersuchen, ob es eine Veränderung in der Selbsteinschätzung hinsichtlich der richtig beantworteten Aufgaben des Baseline Mathematiktest im Vergleich zum nach dem Priming bearbeiteten Mathematiktest gegeben hat, wurde eine Varianzanalyse mit Messwiederholung durchgeführt. Das Ergebnis zeigt einen signifikanten Effekt zwischen den Selbsteinschätzungen von Pre-Priming zum Post-Priming Mathematiktest. Es gab also eine Veränderung in der eigenen Einschätzung. Sowohl weibliche als auch männliche Teilnehmer schätzen beim Pre-Priming Mathematiktest mehr Aufgaben richtig beantwortet zu haben als beim nach dem Priming bearbeiteten Mathematiktest. Tabelle 32 zeigt die durchschnittliche Einschätzung der eigenen Leistung insgesamt sowie für beide Geschlechter getrennt und für die einzelnen Primingbedingungen.

**Tabelle 32: Deskriptive Statistiken der eigenen Einschätzungen bezüglich der erreichten Punktezah
in dem Mathematiktest vor und dem Mathematiktest nach dem Priming**

	Geschlecht	Priming	Mittelwert	St. Abweichung	N
Eigene Einschätzung Pre-Priming Mathematiktest	Weiblich	Stereotypnonkonform	3.689	1.467	61
		Stereotypkonform	3.638	1.398	58
		Gesamt	3.664	1.428	119
	Männliche	Stereotypnonkonform	4.564	1.411	55
		Stereotypkonform	4.726	1.674	51
		Gesamt	4.642	1.538	106
	Gesamt	Stereotypnonkonform	4.103	1.500	116
		Stereotypkonform	4.147	1.620	109
		Gesamt	4.124	1.556	225
Eigene Einschätzung Post-Priming Mathematiktest	Weiblich	Stereotypnonkonform	3.197	1.730	61
		Stereotypkonform	3.224	1.920	58
		Gesamt	3.210	1.817	119
	Männlich	Stereotypnonkonform	3.818	2.212	55
		Stereotypkonform	4.353	2.008	51
		Gesamt	4.075	2.123	106
	Gesamt	Stereotypnonkonform	3.491	1.989	116
		Stereotypkonform	3.752	2.033	109
		Gesamt	3.618	2.010	225

Die Varianzanalyse mit Messwiederholung (Tabelle 32) zeigt ein signifikantes Ergebnis für das Geschlecht: Männer und Frauen schätzen sich also unterschiedlich gut ein. Sowohl weibliche wie auch männliche Teilnehmer schätzen, beim Pre-Priming Mathematiktest mehr Aufgaben richtig beantwortet zu haben, als bei dem Mathematiktest, welchen sie nach dem Priming bearbeiten. Allerdings gibt es keine signifikante Wechselwirkung zwischen dem Geschlecht und der Einschätzung der eigenen Leistung in den Tests. Das Priming hat keinen signifikanten Effekt auf die Einschätzung der eigenen Leistungen. „Selbsteinschätzung“ in Tabelle 33 meint hier, die eigene Einschätzung der in den Mathetests richtig beantworteten Aufgaben. Levene's Test zeigt kein signifikantes Ergebnis für die Selbsteinschätzungen nach dem Baseline Mathematiktest und auch kein signifikantes Ergebnis der angegebenen Selbsteinschätzung für den nach dem Priming vorgegebenen Mathematiktest (Post-Priming Mathematiktest). Das Priming hat keinen signifikanten Effekt auf die Einschätzung der eigenen Leistungen.

Tabelle 33: Ergebnisse der Varianzanalyse mit Messwiederholung bezüglich der Selbsteinschätzung der erreichten Punkte in den Mathematiktests vor bzw. nach dem Priming

Varianzanalyse mit Messwiederholung				
Effekt	F	Sig.	MQ	df
Selbsteinschätzung	19.266	0.000	28.666	1
Geschlecht	21.015	0.000	96.507	1
Priming	0.691	0.407	3.175	1
Geschlecht x Priming	0.790	0.375	3.627	1
Selbsteinschätzung x Priming	0.957	0.329	1.423	1
Selbsteinschätzung x Geschlecht	0.212	0.645	0.316	1
Selbsteinschätzung x Priming x Geschlecht	0.409	0.523	0.609	1
Levene Test	F	Sig	df1	df2
Selbsteinschätzung Pre-Priming Mathematiktest	0.736	0.532	3	221
Selbsteinschätzung Post – Priming Mathematiktest	1.151	0.329	3	221

Mittels einer univariaten Varianzanalyse (Anhang E, Tabelle 46) kann gezeigt werden, dass weibliche und männliche Teilnehmer sich bezüglich ihrer Einschätzung des nach dem Priming bearbeiteten Mathematiktests unterscheiden ($F(3, 224) = 11.08, p = 0.01$), jedoch kann kein signifikanter Effekt des Primings ($F(3, 224) = 1.143, p = 0.286$) nachgewiesen werden.

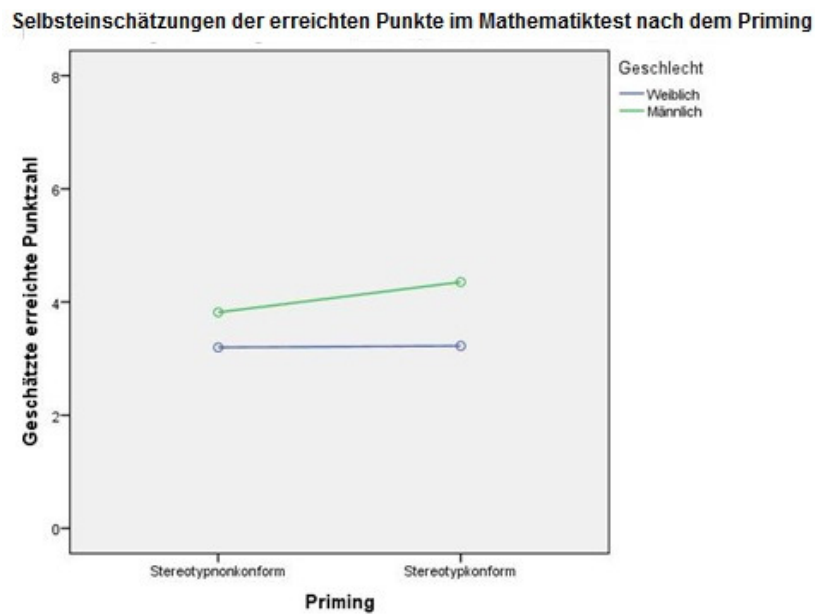


Abbildung 13: Interaktionsdiagramm der Selbsteinschätzungen erreichter Punkte im Mathematiktest nach dem Priming für männliche und weibliche Teilnehmer in den verschiedenen Primingbedingungen

Wie Abbildung 13 zeigt, schätzten sich männliche Teilnehmer signifikant besser ein als weibliche. Außerdem schätzten männliche Teilnehmer in der stereotypkonformen Bedingung ihre erreichte Punktzahl höher ein, als jene in der stereotypkonformen Bedingung, dies ist jedoch nicht signifikant.

3.2.4 Beeinflusst das Priming die Fremdeinschätzung?

Die Teilnehmer wurden nach dem Baseline Mathematiktest (vor dem Priming) danach gefragt, was sie dachten, wie viele Aufgaben andere Mädchen bzw. Burschen ihres Alters richtig hatten. Eine solche Befragung fand ebenfalls nach der Bearbeitung des Mathematiktest, welcher nach dem Priming vorgegeben wurde, statt. Die Teilnehmer konnten Zahlen von 0 bis 8 in die dafür vorgesehen Felder eintragen. Bei den Mathematiktestteilen, welche vor dem Priming vorgegeben wurden, konnten sie Zahlen von 0 bis 4 eingeben. Die nach dem Baseline Mathematiktest (vor dem Priming) eingetragenen Werte aus beiden Terminen wurden addiert, um eine Schätzung für den gesamten Pre-Priming Mathematiktest zu bekommen.

Bei der Frage: „Was denkst du ,wie viele Aufgaben andere Burschen deines Alters richtig hatten?“ zeigt eine Varianzanalyse mit Messwiederholung, dass sowohl Burschen als auch Mädchen schätzten, dass andere Burschen ihres Alters beim Mathematiktest nach dem Priming durchschnittlich weniger Aufgaben richtig hatten, als in dem Mathetest vor dem Priming. Weder das Geschlecht noch die Priming Bedingung zeigen einen signifikanten Effekt. Tabelle 34 zeigt die Ergebnisse der Varianzanalyse mit Messwiederholung für beide Geschlechter und alle Priming Bedingungen bezüglich der Frage, ob sich die Fremdeinschätzung für „andere Burschen ihres Alters“ verändert hat. Für die Erwartungshaltung von weiblichen und männlichen Teilnehmern bezüglich der von Burschen richtig beantworteten Mathematikaufgaben ergab der Levene's Test (Tabelle 34) kein signifikantes Ergebnis für den Pre-Priming Mathematiktest, jedoch sehr wohl ein signifikantes Ergebnis für die Einschätzung des nach dem Priming bearbeiteten Mathematiktest. Die Tabelle 46 mit den Mittelwerten der Teilnehmer bezüglich dieser Frage findet sich im Anhang E.

Tabelle 34: Varianzanalyse mit Messwiederholung der Fremdeinschätzungen der erreichten Punkte in den Mathematiktests von Burschen vor und nach dem Priming

Varianzanalyse mit Messwiederholung				
Effekt	F	Sig.	MQ	df
Fremdeinschätzung	8.324	0.004	10.715	1
Geschlecht	2.953	0.087	2.953	1
Priming	3.525	0.062	3.525	1
Geschlecht x Priming	1.897	0.172	1.897	1
Fremdeinschätzung x Priming	0.001	0.980	0.001	1
Fremdeinschätzung x Geschlecht	0.197	0.658	0.253	1
Fremdeinschätzung x Priming x Geschlecht	0.183	0.669	0.236	1
Levene Test	F	Sig	df1	df2
Fremdeinschätzung Pre-Priming Mathematiktest	1.397	0.245	3	221
Fremdeinschätzung Post – Priming Mathematiktest	6.113	0.001	3	221

Tabelle 35 zeigt das Ergebnis einer Varianzanalyse mit Messwiederholung für die Frage: „Was denkst du, wie viele Aufgaben Mädchen deines Alters richtig hatten?“. Es zeigte sich, dass sich bezüglich der von allen Schülern geschätzten erreichten Punkte von Mädchen in den Mathetests vor bzw. nach dem Priming ein

signifikanter Effekt für das Geschlecht und für die angegebene Punkteinschätzung vor bzw. nach dem Priming findet. Levene's Test weist kein signifikantes Ergebnis für die Fremdeinschätzung von Mädchen erreichten Punkte beim Baseline Mathematiktest (Pre-Priming Mathematiktest), jedoch ein signifikantes Ergebnis für die Fremdeinschätzung von Mädchen erreichten Punkte für den nach dem Priming bearbeiteten Mathematiktest auf.

Tabelle 35: Varianzanalyse mit Messwiederholung der Fremdeinschätzungen der erreichten Punkte in den Mathematiktests von Mädchen vor und nach dem Priming

Varianzanalyse mit Messwiederholung				
Effekt	F	Sig.	MQ	Df
Fremdeinschätzung	13,610	0.000	14.763	1
Geschlecht	5.973	0.015	5.973	1
Priming	0.809	0.181	0.181	1
Geschlecht x Priming	7.073	0.623	0.243	1
Fremdeinschätzung x Priming	0.782	0.378	0.848	1
Fremdeinschätzung x Geschlecht	0.197	0.008	7.672	1
Fremdeinschätzung x Priming x Geschlecht	1.169	0.281	1.268	1
Levene Test	F	Sig	df1	df2
Fremdeinschätzung Pre- Priming Mathematiktest	2.405	0.068	3	221
Fremdeinschätzung Post-Priming Mathematiktest	7.064	0.000	3	221

Betrachtet man die Mittelwerte der weiblichen und männlichen Teilnehmer bezüglich der Einschätzung der von Mädchen erreichten Punkte in den Mathematiktests (Tabelle 36) genauer, zeigt sich, dass weibliche Teilnehmer ihre Schätzung kaum verändern. So schätzen weibliche Teilnehmer, dass andere weibliche Teilnehmer durchschnittlich $M = 4.95$ ($SD = 1.25$) Aufgaben bei dem Mathematiktest vor dem Priming (Baseline) richtig haben und $M = 4.88$ ($SD = 1.47$) bei dem Mathematiktest nach dem Priming. Männliche Teilnehmer kommen auf einen Mittelwert von $M = 4.76$ ($SD = 1.66$) vor dem Priming und $M = 4.13$ ($SD = 2.17$) nach dem Priming.

Tabelle 36: Mittelwerte, Standardabweichungen und Anzahl der Einschätzungen von Mädchen richtig gelösten Aufgaben in den Mathematiktests vor und nach dem Priming

	Geschlecht	Priming	Mittelwert	St. Abweichung	N
Fremdeinschätzung Mädchen Pre-Priming Mathematiktest	Weiblich	Stereotypnonkonform	4.967	1.251	61
		Stereotypkonform	5.000	1.298	58
		Gesamt	4.983	1.269	119
	Männliche	Stereotypnonkonform	4.946	1.682	55
		Stereotypkonform	4.569	1.628	51
		Gesamt	4.769	1.659	106
	Gesamt	Stereotypnonkonform	4.957	1.465	116
		Stereotypkonform	4.798	1.471	109
		Gesamt	4.880	1.467	225
Fremdeinschätzung Mädchen Post-Priming Mathematiktest	Weiblich	Stereotypnonkonform	4.885	1.496	61
		Stereotypkonform	4.879	1.476	58
		Gesamt	4.882	1.480	119
	Männlich	Stereotypnonkonform	4.127	2.310	55
		Stereotypkonform	4.137	2.030	51
		Gesamt	4.132	2.169	106
	Gesamt	Stereotypnonkonform	4.526	1.954	116
		Stereotypkonform	4.532	1.788	109
		Gesamt	4.528	1.871	225

Ein t-Test für abhängige Stichproben zeigt, dass der Mittelwertunterschied bezüglich der Einschätzung der richtig beantworteten Mathematikaufgaben (vor bzw. nach dem Priming) von Mädchen, bei männlichen Teilnehmern signifikant ($t(105) = 3.724$, $p = 0.000$) ausfällt. Männliche Versuchspersonen schätzten, dass weibliche Teilnehmer beim Mathematiktest nach dem Priming signifikant weniger Aufgaben richtig hatten als im Mathematiktest vor dem Priming. Eine univariate Varianzanalyse (Anhang E, Tabelle 47) zeigt außerdem, dass männliche und weibliche Teilnehmer sich hinsichtlich der Einschätzung der im Post-Priming, Mathematiktest von Mädchen richtig bearbeiteten Aufgaben signifikant ($F(3, 224) = 3.952$, $p = 0.002$) unterscheiden. Abbildung 13 zeigt ein Interaktionsdiagramm der Einschätzung durchschnittlich richtig gelöster Aufgaben im Mathematiktest nach dem Priming für männliche und weibliche Teilnehmer in den verschiedenen Primingbedingungen, hier zeigt sich nochmals, dass die Primingbedingungen einen Einfluss auf die Fremdeinschätzungen zu haben scheinen.

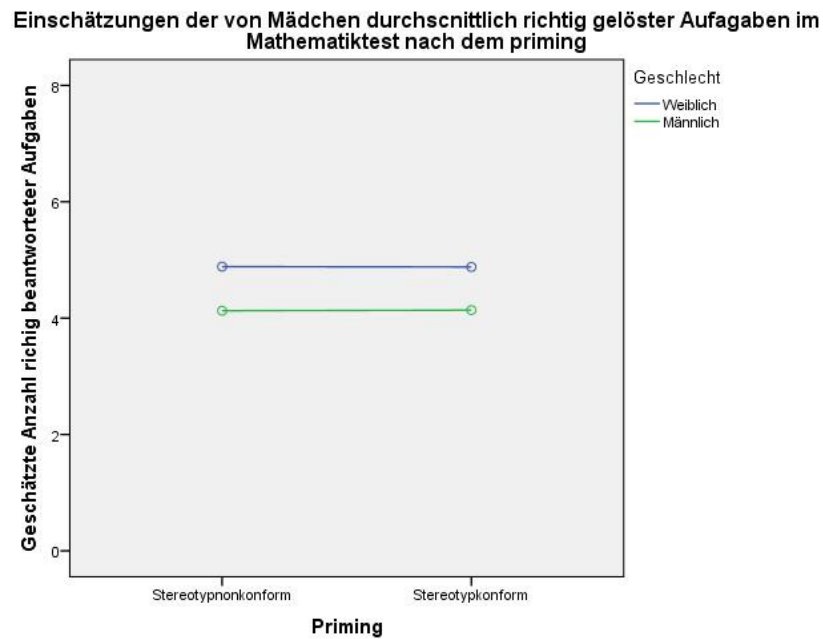


Abbildung 14: Interaktionsdiagramm der Einschätzung durchschnittlich richtig gelöster Aufgaben im Mathematiktest nach dem Priming für männliche und weibliche Teilnehmer in den verschiedenen Primingbedingungen

3.2.5 Wenn die Teilnehmer eine geringe Motivation haben, hat das Priming keinen Einfluss auf die Mathematiktestwerte der Teilnehmer

Wenn man die Motivation der Teilnehmer zum Termin betrachtet, in dem sie den Mathematiktest nach dem Priming bearbeiteten, sieht man, dass eine Tendenz zur Mitte hin besteht. Die Motivation konnte auf einer Skala von 1 bis 6 eingeschätzt werden. Abbildung 15 zeigt Näheres zu der Motivation der Teilnehmer, bei dem Mathematiktest nach dem Priming. Die Teilnehmer gaben größtenteils eine durchschnittliche Motivation an, daher ist es unwahrscheinlich, dass dies als mögliche Ursache für das Fehlen des erwarteten Effekts des Primings in Frage kommt. Auch der Ausschluss der Teilnehmer mit einer sehr niedrigen Motivation (1 oder 2) brachte keine signifikanten Ergebnisse in einer Varianzanalyse (Siehe Anhang E, Tabelle 41, 42 und 43).

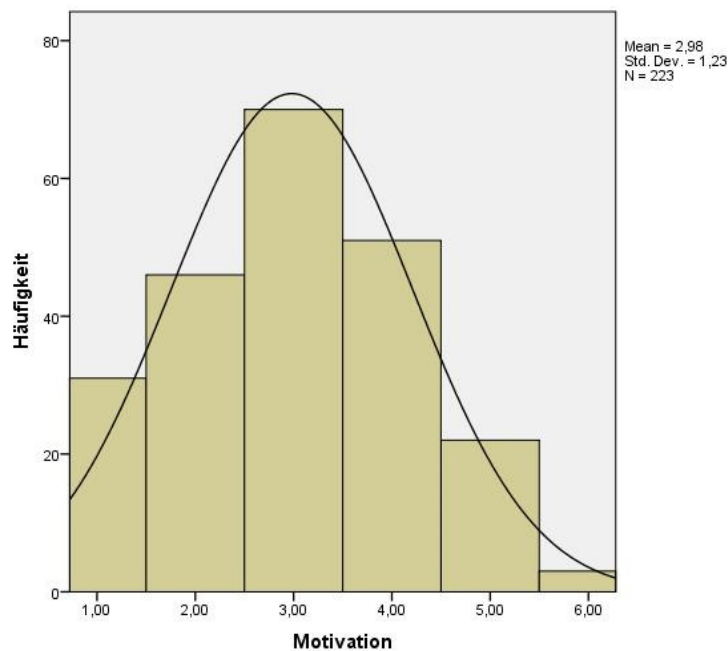


Abbildung 15: Histogramm der eigenen Einschätzung der Motivation der Teilnehmer

3.2.6 Hat die Primingmethode einen Einfluss auf die Stärke des Stereotype Threat, oder Stereotype Lift Effekt?

Da sich weder bei dem direkten Priming, noch bei der subtileren Variante (Reischer, 2015) ein signifikanter Effekt der jeweiligen Primingbedingungen nachweisen ließ, kann diese Fragestellung an dieser Stelle nicht beantwortet werden.

3.3 Weitere Ergebnisse

3.3.1 Stereotype Einstellungen bezüglich Mathematik vor der Bearbeitung von Aufgaben

Da vorhandene Vorurteile und Stereotype Endorsement wichtige Mediatoren für das Auftreten von Stereotype Lift und Threat Effekten sind, wurde versucht diese mit in die Untersuchung einzubinden. Um einzuschätzen, ob mögliche Priming-Effekte durch bereits vorhandene Stereotypeinstellungen der Teilnehmer bezüglich der Leistungen von Mädchen und Burschen in Mathematik bestehen, wurde dies ebenfalls in den demographischen Daten erfragt. Die Ergebnisse zeigen, dass weder die weiblichen noch männlichen Teilnehmer insgesamt eine

besonders stereotype Einstellung bezüglich der mathematischen Leistungen von ihren Mitschülern hatten.

Tabelle 37: Mediane und Anzahl der Einschätzungen weiblichen und männlichen Teilnehmer bezüglich mathematischer Fähigkeiten von dem Pre-Priming Mathematiktest und dem Post-Priming Mathematiktest

Geschlecht		Einschätzung der Mathematikfähigkeit von Mädchen	Einschätzung der Mathematikfähigkeit von Burschen
Weiblich			
Weiblich	N	119	119
	Median	3	4
Männlich			
Männlich	N	106	106
	Median	4	3

Weibliche Teilnehmer schätzten mit einem Median von 4, dass Burschen besser abschnitten als Mädchen (Median: 3) und männliche Teilnehmer schätzten, dass Mädchen besser als Burschen abschnitten, jedoch ist dies bei den Burschen nicht signifikant. Tabelle 37 zeigt die Mediane der Untersuchung der Einschätzungen der mathematischen Fähigkeiten von weiblichen und männlichen Personen vor dem Matrizentest, Pre-Priming Mathematiktest und dem Post-Priming. Ein U-Test für unverbundene Stichproben zeigt, dass sich die Geschlechter bezüglich ihrer Einschätzung dessen wie gut Mädchen in Mathematik sind, nicht signifikant unterscheiden ($z = -1.393$, $U = 5647$, $p = 0.164$). Bei der Einschätzung der Burschen unterscheiden sich die Geschlechter Einschätzungen jedoch ($z = -3.307$, $U = 4736.5$, $p = 0.001$). Weibliche Teilnehmer schätzen die mathematischen Fähigkeiten von Burschen signifikant höher ein, als die männlichen Teilnehmer.

3.3.2 Geschlechtsunterschiede bezüglich der Arbeitshaltungen

Geschlechterübergreifend zeigt sich bei allen Arbeitshaltungen eine Tendenz zur Mitte (Median 3 oder 4 bei allen Arbeitshaltungen). Tabelle 38 zeigt die Median, sowie die Quartile der verschiedenen Arbeitshaltungen getrennt für beide Geschlechter. Untersucht wurden 223 Personen, welche beide Testteile ganz bearbeiteten. Zwei Person wurde ausgeschlossen, da diese die Arbeitshaltungen (vermutlich wegen eines abgestürzten PCs) nicht bearbeitet hatten. Die Arbeitshaltungen beziehen sich jeweils auf den Termin, an dem die

Versuchspersonen den Mathematiktest bearbeiteten, bei dem mittels eines Artikels geprint wurde. Die Teilnehmer konnten die verschiedenen Bereiche auf einer Skala von 1 bis 6 einschätzen. Im Anhang D finden sich die Histogramme der einzelnen Fähigkeiten. Da es sich um Rangdaten handelt werden U-tests eingesetzt um zu untersuchen, ob weibliche und männliche Teilnehmer unterschiedlich motiviert waren, sich bezüglich der empfunden Schwierigkeit der Mathematikaufgaben, der aufgewandten Mühe (welche sie sich bei der Bearbeitung gaben), der empfundenen Geschwindigkeit mit der sie arbeiteten, des angegebenen Interesses, sowie der gefühlten Forderung unterscheiden.

Tabelle 38: Median und Quartile der Arbeitshaltungen bei dem Termin an dem die Versuchspersonen mittels Artikel geprint wurden

	Geschlecht	Median	Quartile	
			25	75
Konzentration	Weiblich	3	3	4
	Männlich	3	2	4
Geschwindigkeit	Weiblich	4	3	4
	Männlich	4	3	5
Interesse	Weiblich	3	2	4
	Männlich	4	3	4
Forderung	Weiblich	3	3	4
	Männlich	3	2	4
Bemühung	Weiblich	4	3	4
	Männlich	3	3	4
Schwierigkeit	Weiblich	4	3	5
	Männlich	4	3	4
Motivation	Weiblich	3	2	4
	Männlich	3	2	4

Bei der Betrachtung der Ergebnisse des U-Tests, zeigten sich Geschlechtsunterschiede in der Einschätzung der angegebenen Geschwindigkeit, des Interesses, der Forderung und der Schwierigkeit. Männliche Versuchsteilnehmer gaben an, schneller zu arbeiten und ein höheres Interesse an den Aufgaben der Studie zu haben. Männliche Teilnehmer fühlten sich auch weniger gefordert und schätzten sie die Schwierigkeit der Mathematikaufgaben niedriger ein, als die weiblichen Teilnehmer.

Tabelle 39: Ergebnisse des U-Tests bezüglich der Arbeitshaltungen

	Geschlecht	Mittlerer Rang	Summe der Ränge	U	Z	Sig
Konzentration	Weiblich	116.903	13794.50	561.,500	-1.233	0.217
	Männlich	106.491	11181.50			
Geschwindigkeit	Weiblich	101.886	12022.50	5001.500	-2.574	0.010
	Männlich	123.367	12953.50			
Interesse	Weiblich	102.229	12063.00	5042.000	-2.479	0.013
	Männlich	122.981	12913.00			
Forderung	Weiblich	124.352	14673.50	4737.500	-3.209	0.001
	Männlich	98.120	10302.50			
Bemühung	Weiblich	118.691	14005.50	5405.500	-1.697	0.090
	Männlich	104.481	10970.50			
Schwierigkeit	Weiblich	124.678	14712.00	4699.000	-3.212	0.001
	Männlich	97.752	10264.00			
Motivation	Weiblich	115.364	13613.00	5798.000	-0.849	0.396
	Männlich	108.219	11363.00			

3.3.3 Zusammenhänge zwischen den Arbeitshaltungen

Relevante signifikante Zusammenhänge zwischen den Arbeitshaltungseinstellungen (bei dem Termin an welchem auch der Post-Priming Mathematiktest bearbeitet wurde) zeigten sich zwischen Konzentration und Interesse ($r=0.373$), Konzentration und Bemühung ($r=0.620$), Konzentration und Motivation ($r=0.523$), Konzentration und Geschwindigkeit ($r=0.322$), Geschwindigkeit und empfundener Schwierigkeit ($r=-0.317$), Interesse und Bemühung ($r=0.533$), Interesse und Motivation ($r=0.461$), Forderung und empfundener Schwierigkeit der Mathematikaufgaben ($r=0.449$) und schließlich, zwischen Bemühung und Motivation ($r=0.475$). Personen welche angaben sehr konzentriert zu sein, waren daher auch eher bemüht, hatten mehr Interesse und waren motivierter bei der Bearbeitung der Aufgaben. Je interessierter die Teilnehmer waren, desto bemühter und motivierter waren sie. Je schwieriger die Teilnehmer die Aufgaben bewerteten, desto langsamer arbeiteten sie. Personen welche sich gefordert fühlten, empfanden auch die Schwierigkeit der Aufgaben höher. Schließlich zeigte sich noch, dass bemühte Personen, angaben motivierter zu sein. Ein weiterer signifikanter Zusammenhang zeigte sich zwischen der Einschätzung der

Schwierigkeit der Mathematikaufgaben und der Motivation ($r = -0.296$). Je schwieriger die Teilnehmer die Mathematikaufgaben einschätzten, desto weniger motiviert waren sie bei deren Bearbeitung.

4 Diskussion

4.1 Ergebnisse der Untersuchung

Die Reliabilität unserer Tests liegt bei dem Baseline Test bei 0,41 und dem Mathematiktest welcher nach dem Priming vorgegeben wurde bei 0,57, welches für nur 8 Aufgaben kein schlechtes Ergebnis ist. Zwei der Aufgaben schienen den Teilnehmern besondere Schwierigkeiten zu bereiten, diese sollte man daher aus einer Wiederholungsstudie ausschließen. Es handelt sich dabei um Aufgabe 5 und 7. Sowohl im Pre-Priming, wie auch im Post-Priming Mathematiktest, hatten diese beiden Aufgaben die geringste Lösungswahrscheinlichkeit. Außerdem wurde eine der Mathematikaufgaben (Aufgaben 8) häufiger von den männlichen, als von weiblichen Teilnehmern gelöst. Möglicherweise hängt dies mit den Lösungsstrategien, oder der Natur der Aufgabe zusammen. Aufgrund der Tatsache, dass nur wenige Teilnehmer ihren Rechenweg aufschrieben, ist es leider in dieser Studie nicht möglich nachzuvollziehen, warum diese Aufgabe so unterschiedlich von den männlichen bzw. weiblichen Teilnehmern bearbeitet wurde. Die Aufgabenschwierigkeit schwankte zwischen 7,5 (für Aufgabe sieben) und 86 (für Aufgabe eins) bei dem Baseline Mathematiktest und 5 (für Aufgabe sieben) und 73,5 (Bei Aufgabe eins) bei dem nach dem Priming vorgegebenen Mathematiktest. Dies lässt eigentlich auf eine angemessene Schwierigkeit der Tests schließen, dennoch zeigte sich leider kein unterschiedlicher Effekt der verschiedenen Bedingungen auf die Leistungen. Da die Aufgaben in beiden Tests eine ähnliche Lösungswahrscheinlichkeit aufweisen, sind diese zwar gut parallelisiert, jedoch zeigt sich hier nochmals, dass unsere Manipulationsversuche wenig bis keinen Einfluss auf die Lösungs-wahrscheinlichkeiten der Aufgaben hatten.

Die Betrachtung der relevanten Teilnehmer bestätigt, dass es keinen signifikanten Geschlechtsunterschied in der Grundfähigkeit Mathematik gibt. Es zeigt sich ein ähnlichen Bild wie bei der Untersuchung von Hyde et al. (2008). Allerdings sieht dieses Ergebnis anders aus wenn bestimmte Versuchsteilnehmer ausgeschlossen werden bzw. berücksichtigt werden. Für nähere Ergebnisse dazu

möchte ich auf die Diplomarbeit meiner Studienkollegin Regina Reischer verweisen. Sie schätzte bei den Teilnehmern, welche nur die Hälfte des Baseline Tests bearbeitet hatten (daher bei einem Termin fehlten), die fehlenden Werte unter zu Grundlegung ihrer Leistung in dem bearbeiteten Teil der Baseline Untersuchung, mittels linearer Regression.

Unerwarteter Weise verschlechterten sich alle Versuchspersonen nach dem Priming, allerdings verschlechterten sich die männlichen Teilnehmer in beiden Versuchsbedingungen signifikant, bezüglich ihrer Leistungen in dem nach dem Priming absolvierten Mathematiktest (Post-Priming), im Vergleich zum vor dem Priming bearbeiteten Mathematiktest (Pre-Priming). Möglicherweise liegt das daran, dass weibliche Teilnehmer bemühter waren bei der Bearbeitung des Tests, jedoch ist der angegebene Unterschied, hinsichtlich der Bemühung, nicht signifikant. Männliche Teilnehmer gaben in den Arbeitshaltungen außerdem an, schneller zu arbeiten, ein höheres Interesse an den Aufgaben der Studie zu haben und sich auch weniger gefordert gefühlt zu haben. Zudem schätzten sie die Schwierigkeit niedriger ein, als die weiblichen Teilnehmer. Die Motivation der Teilnehmer war größtenteils durchschnittlich. Zwar gaben auch relativ wenige Personen an, sehr motiviert zu sein, aber da unsere Studie ihnen auch keine Extrapunkte oder ähnliches brachte, ist dies nicht weiter verwunderlich. Hätten die Teilnehmer vermehrt angegeben sehr unmotiviert zu sein, hätte man annehmen könnten, dass unser Priming deswegen keinen Effekt gehabt haben könnte.

Sehr voreingenommene Personen neigen eher dazu, Vergleiche nach unten anzustreben (Wills, 1981) und werden daher mit einer höheren ,Wahrscheinlichkeit von den Vorteilen von Stereotype Lift Effekten beeinflusst (Walton & Cohan, 2003). In unserer Stichprobe waren nicht nur die weiblichen Teilnehmer voreingenommen, in Bezug auf die mathematischen Fähigkeiten von Frauen und Männern. Im Rahmen der demographischen Daten wurden die weiblichen und männlichen Teilnehmer darüber befragt, wie gut sie die mathematischen Fähigkeiten von Mädchen und Burschen (ihres Alters) einschätzten. Bei den weiblichen Versuchspersonen zeigte sich, dass sie die mathematischen Fähigkeiten von Mädchen und Burschen signifikant unterschiedlich einschätzten. Weibliche Teilnehmer schätzten, dass Burschen

besser in mathematischen Bereichen abschneiden als Mädchen, welches genau dem gängigen Stereotyp entspricht. Bei männlichen Teilnehmern konnte dieser Trend nicht gefunden werden. Auch die Fremdeinschätzungen (Einschätzungen der Leistungen anderer Personen in den vorgegebenen Mathematiktests, des eigenen und anderen Geschlechts) wurden nicht durch das Priming beeinflusst. Mädchen veränderten ihre Einschätzung der von anderen Mädchen richtig beantworteten Aufgaben vom ersten (Pre-Priming Mathematiktest) zum zweiten Mathematiktest durchschnittlich kaum. Burschen dagegen schätzten, dass Mädchen beim Post-Priming Mathematiktest signifikant schlechter abschnitten, als bei jenem Mathematiktest vor dem Priming. Bei der Einschätzung von den Leistungen von Burschen zeigten sich keine signifikanten Effekte.

Cadinu, Maass, Frigerio, Impagliazzo und Latinotti (2003) zeigten, dass weibliche Personen denen mitgeteilt wurde, sie würden besser abschneiden als Männer, tatsächlich besser abschnitten als jene denen mitgeteilt wurde, dass Männer besser abschneiden. Dies zeigte sich jedoch nur bei Personen mit einer hohen Domänenidentifikation. Wir fragten nach der persönlichen Wichtigkeit von Mathematik, konnten aber keinen Effekt dieser Einschätzung auf die Leistungen nachweisen. Auch wenn nur Personen in die Berechnung einbezogen wurden, welche eine hohe persönliche Wichtigkeit angegeben hatten, konnte kein signifikanter Effekt des Primings nachgewiesen werden.

4.2 Kritik und Ausblick

In dieser Studie sollte unter anderem die Wirksamkeit eines Primings mittels einer Baseline getestet werden. Die Ausarbeitung der Methode unter der Voraussetzung, nur limitierte Zeit zur Verfügung zu haben und zwei Priming-Methoden (zum späteren Vergleich) zu verwenden, war ein schwieriges Unterfangen. Da allerdings weder bei der einen, noch der anderen Priming-Methode ein signifikanter Effekt von Priming festzustellen war, erübrigt sich der geplante Vergleich der Priming-Methoden. Der Ablauf der Studie, war ausführlich geplant und umgesetzt, um einen problemfreien Ablauf sicherzustellen. Die Umsetzung des geplanten Ablaufs verlief relativ problemfrei. Schwierigkeiten ergaben sich unter anderem mit der Schulbehörde. Wir hatten insgesamt vier Schulstunden für

die Durchführung der Studie veranschlagt, welches nur schwer in den Lehrplan der Schüler zu integrieren war. Nach unserer ersten Testung stellten wir jedoch fest, dass die Zeiten die wir für die verschiedenen Teile der Testung vorgesehen hatten, nicht mit der eigentlichen Bearbeitungszeit überein stimmten. Wir hatten im ersten Entwurf 15 Minuten für die Bearbeitung des nach dem Priming vorgegebenen Mathematiktest vorgesehen. Dieser Zeitraum stellte sich jedoch für die meisten Teilnehmer unserer ersten Testklasse als viel zu kurz dar, um alle Aufgaben zu bearbeiten. Außerdem zeigte sich, dass die Schüler viel schneller als wir es erwarteten mit der Bearbeitung der demographischen Daten, wie auch dem Lesen des Artikels fertig wurden. Dies gab uns die Möglichkeit die Zeit für die Bearbeitung des Baseline Mathematiktest (Pre-Priming) von sieben auf 13 Minuten und die Bearbeitungszeit des nach dem Priming (Post-Priming) vorgegebenen Mathematiktest auf 25 Minuten zu erhöhen. Mit einer gezielten Vorstudie hätten diese Probleme vermieden werden können. Die Mathematikaufgaben selbst waren so gewählt, dass sie einem Mathematiktest welcher auch in dieser Form in einer Unterrichtsstunde vorgegeben werden würden sehr nahe kamen. Es sollte so sichergestellt werden, dass die Schüler diese Art Aufgaben in jeden Fall auch bereits in der Schule behandelt hatten und ihren Fähigkeiten entsprechen. Um sicherzugehen dass die Mathematikaufgaben alltagsnahe für die Schüler sind wurden sie unter Absprache mit einigen Mathematiklehrern, welche die getestete Schulstufe auch unterrichten entwickelt und angepasst.

Die technischen Probleme, welche auftraten, waren zwar meist lösbar, jedoch kostete dies immer wieder Zeit und verfälschte so auch die Zeit, welche die einzelnen Teilnehmer für die Bearbeitung der Studie brauchten. Wenn der Bildschirm eines Teilnehmers aktualisiert werden musste, wurden die bereits eingetragenen Lösungen gelöscht und mussten neu eingetragen werden. Wenn eine Aktualisierung der Seite während der Bearbeitung des Mathematiktests notwendig wurde, mussten die Schüler ihre Lösungen neu eintragen. Zwar wurden die Teilnehmer darauf aufmerksam gemacht, dass sie ihre Antworten auf dem dazu vorgesehen Zettel aufschreiben sollten, jedoch kamen viele Teilnehmer dieser Bitte nicht nach. Leider nahmen auch viele Schüler die Möglichkeit nicht

wahr, ihre Rechenwege und Antworten, auf dem für sie bereitgestellten (und später eingesammelten) Zettel, aufzuschreiben. Einige unserer Teilnehmer hatten Schwierigkeiten damit, ihren eigenen Code zu erstellen und verwendeten stattdessen den Beispielcode. Dennoch war es uns möglich, anhand der Handschrift und der passenden Daten, die meisten Zettel der beiden Termine einander zuzuordnen. Ein weiteres Problem ergab sich mit dem, in die Umfragesoftware LimeSurvey eingearbeiteten, Taschenrechner. Zwar verstanden alle Teilnehmer, wie dieser zu benutzen sei, jedoch vergaßen leider einige Teilnehmer ihre Lösungen in die dafür vorgesehenen Antwortfelder einzutragen, oder trugen ihre Antworten direkt in das Taschenrechnerfeld ein.

Falls man diese Studie replizieren wollen sollte, würde ich dazu raten, sich nur für eine Priming-Methode zu entscheiden. So könnte man an einem Termin ohne Priming arbeiten und am Zweiten mit der gewählten Priming-Methode. Ebenfalls könnte man auf diesem Wege feststellen, ob der in dieser Studie nicht signifikante Primingeffekt, möglicherweise auf die Ermüdung der Schüler, oder auf die Halbierung der Testmethode zurückzuführen ist. Dafür würde sprechen, dass alle Teilnehmer in jeder Gruppe im zweiten Test (nach dem Priming) schlechtere Leistungen brachten, als im ersten Baseline Test. Natürlich hätten wir uns auch für einen Multiple Choice Test entscheiden können, jedoch hatten wir die Befürchtung, dass die Teilnehmer einfach raten würden. Daher würde ich bei einer Wiederholung dieser Studie sicherlich wieder zu Textaufgaben ohne MC Format raten.

Wieso das Priming keinen signifikanten Effekt auf die Schüler hatte, könnte verschiedene Gründe gehabt haben. Zu der Zeit der Testung wurde über die Ergebnisse der aktuellsten PISA Studie in den Medien berichtet, diese unterstützten den typischen Stereotyp, dass Männer besser in mathematischen Bereichen seien. Jedoch hätte gerade dieses dazu führen sollten, dass sich der erwartete Effekt bei den weiblichen Teilnehmern einstellt. Möglicherweise lag das Problem auch an der nicht adäquaten, oder typischen Präsentation des Artikels im Limesurvey. Jedoch hielt ein Großteil der Versuchsteilnehmer (unabhängig von der Versuchsbedingung und daher auch vom Inhalt) diesen für relativ glaubwürdig. Allerdings zeigen auch Studien von Huguet und Régner (2009), dass gegenteilige

Ansichten, hinsichtlich des gängigen Stereotyps bezüglich der mathematischen Fähigkeiten von weiblichen Personen, keine Mediatorvariable für das Auftreten eines Stereotype Threat Effekts darstellen. Anders als bei Appel et al. (2011) wurden die Fragen zum Text nicht direkt nach dem Text, sondern erst nach dem Mathematiktest gestellt. Dies sollte dazu führen, dass die Schüler sich intensiv mit den Inhalten des Textes befassen. Möglicherweise versuchten die Schüler auch, während der Bearbeitung des Post-Priming Mathematiktest die Inhalte des Artikels im Kopf zu behalten, was einen Teil ihrer Aufmerksamkeit von der Bearbeitung der Aufgaben abziehen würde. Dies würde zumindest eine zusätzliche Erklärung dafür liefern, wieso alle Versuchsgruppen in dem Post Priming Mathematiktest schlechter abschnitten.

Nach dem Ende der Testungen wurden die Schüler über den Sinn und Zweck unserer Studie aufgeklärt. Bei dieser Aufklärung fragten wir auch, ob sich die Teilnehmer von den Priming-Methoden beeinflusst gefühlt hatten. Allgemein gaben nur wenige Personen an, sich von dem Priming beeinflusst gefühlt zu haben. Leider liegen uns keine teststatistischen Daten darüber vor, welche Personen sich beeinflusst gefühlt hatten und welche nicht. Jedoch sagte uns eine männliche Testperson nach dem stereotypnonkonformen Priming (Frauen erbringen bessere Leistungen), er hätte sich durch diese Information herausgefordert gefühlt und hätte sich dann auch mehr Mühe gegeben. Johnson et al. (2012) schlugen bei Ihrer Untersuchung die Erklärung vor, dass männliche Personen sich möglicherweise durch solch eine Information herausgefordert fühlen und daher, wie in ihrer Studie gesehen, ihre Leistungen entsprechend anpassen. Es wäre sicherlich eine interessante Forschungsfrage für folgende Untersuchungen, ob auch andere Personen auf ähnliche Information so reagieren und woran diese Steigerung der Motivation liegen könnte. So konnten mit dieser Untersuchung zwar keine Primingeffekte nachgewiesen werden, oder ihre Ursachen weiter untersucht werden, jedoch ergeben sich für zukünftige Studien einige interessante Ideen und Untersuchungsempfehlungen.

Literatur

- Abrams, D., Crisp, R. J., Marques, S., Fagg, E., Bedford, L. & Provias, D. (2008). Threat inoculation: Experienced and imagined intergenerational contact prevents stereotype threat effects on older people's math performance. *Psychology and Aging*, 23, 934–939.
- Alfermann, D. (1996). *Geschlechtsrollen und geschlechtstypisches Verhalten*. Stuttgart: Kohlhammer.
- Ambady, N., Shih, M., Kim, A. & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychology Science*, 12 (5), 385-390.
- Appel, M., Kronberger, N. & Aronson, J. (2011). Stereotype threat impairs ability building: Effects on test preparation among woman in science and technology. *European Journal of social psychology*, 41, 904-913.
- Aronson, J. (2004). The threat of stereotype. *Educational Leadership*, 62, 14-19.
- 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.
- Ashmore, R. D. & Del Boca F. K. (1979). Sex Stereotypes and Implicit Personality Theory: Toward a Cognitive-social Psychological Conceptualization. *Sex Roles*, 5, 219-248.
- Bargh J. A., Chen, M. & Burrows, L. (1996). Automaticity of social behavior. Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230-244.
- Bargh, J. A. & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, 43, 437-449.

- Bergmann, W. (2001). *Was sind Vorurteile: Vorurteile, Stereotype, Feindbilder, Informationen zur politischen Bildung*. Heft 271, hrsg. v. Bundeszentrale für politische Bildung, Bonn, 3-9. (abgerufen am 13.06.2014, <http://www.bpb.de/izpb/9680/was-sind-vorurteile>).
- Berk, L. E. (2011). *Entwicklungspsychologie*. (5., aktualisierte Aufl.). München: Pearson Studium.
- 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.
- Chalabaev, A., Stone, J., Sarrazin, P., & Croizet, J. C. (2008). Investigating physiological and self-reported mediators of stereotype lift effects on a motor task. *Basic and Applied Social Psychology*, 30, 18-26.
- Chatard, A., Selimbegovic, L., Konan, N. P., & Mugny, G. (2008). Performance boosts in the classroom: Stereotype endorsement and prejudice moderate stereotype lift. *Journal of Experimental Social Psychology*, 44, 1421-1424.
- Croizet, J.-C., Després, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P. & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, 30, 721-731.
- Dar-Nimrod, I. & Heine, S. (2006). Exposure to Scientific Theories Affects Women's Math Performance. *Science*, 314, 435-435.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.
- Dijksterhuis, A., & van Knippenberg, A. (1996). The knife that cuts both ways: Facilitated and inhibited access to traits as a result of stereotype activation. *Journal of Experimental Social Psychology*, 32, 271-288.

- Dröge, F. W. (1967). *Publizistik und Vorurteil*. Münster.
- Eckes, T. (2010). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen
In R. Becker und B. Kortendiek (Hrsg.), *Handbuch Frauen und Geschlechterforschung: Theorie, Methoden, Empirie* (3. Aufl., 178-189).
Wiesbaden: Verlag für Sozialwissenschaften.
- Fazio, R. H.; Sanbonmatsu, D. M., Powell, M. C., Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229-238.
- Good, C., Aronson, J. & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, 29, 17-28.
- Harkins, S. (2006). Mere effort as the mediator of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, 91, 436 – 455.
- Herkner, W. (2003). *Lehrbuch: Sozialpsychologie*, Bern: Hans Huber Verlag
- Huguet, P. & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, 99, 545-560.
- Huguet, P. & Régner, I. (2009). Counter-stereotypic beliefs in math do not protect girls from stereotype threat. *Journal of Experimental Social Psychology*, 45, 1024-1027.
- Hyde, J. S., Fennema, E., Lamon, S. (1990). Gender differences in mathematics performance. A meta-analysis. *Psychological Bulletin*, 107, (2), 139-155.
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321, 494-495.

- Ihme, T. & Mauch, M. (2007). Werbung als implizite Aktivierungsquelle von Geschlechterstereotypen und ihr Einfluss auf Mathematikleistungen sowie auf das Computerwissen bei Mädchen und Jungen. *Empirische Pädagogik*, 21, 291-305.
- Jamieson, J. P. & Harkins, S. G. (2007). Mere Effort and Stereotype Threat Performance Effects. *Journal of Personality and Social Psychology*, 93, 4, 544–564.
- Johnson, H. J., Barnard-Brak, L., Saxon, T. F., & Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education*, 80, 137-149.
- Keller, J. (2002). Blatant Stereotype Threat and Women's Math Performance: Self-Handicapping as a Strategic Means to Cope with Obstrusive Negative Performance Expectations. *Sex Roles*, 47, (3/4), 193-198.
- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' 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.
- Lang, X. (2004). *Der Einfluss von Stereotyp-Priming auf die Raumvorstellungsleistung von Männern und Frauen*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Lee, A. M. I. (2014). *How No Child is left behind affects your child*. <https://www.understood.org/en/school-learning/your-childs-rights/basics-about-childs-rights/how-no-child-left-behind-affects-your-child> (abgerufen am 22.06.2016).
- Lewis, M & Weinraub, M., (1979). Origins of early sex role development. *Sex Roles*, 5, (2), 135- 153.

- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C., (2010). New trends in gender and mathematics performance: a meta-analysis, *Psychological Bulletin*, 136, (6), 1123–1135.
- Martin, C.L., Wood, C.H. & Little, J.K. (1990). The development of gender stereotype components. *Child Development*, 61, 1891-1904.
- Martiny, S. E. & Götz, T. (2011). Stereotype Threat in Lern- und Leistungssituationen: Theoretische Ansätze, empirische Befunde und praktische Implikationen. In M. Dresel & L. Lämmle (Hrsg.), *Motivation, Selbstregulation und Leistungsexzellenz* (S. 153-177). Münster: LIT.
- Marx M.D., & Stapel, D. A. (2006). Distinguishing stereotype threat from priming effects: On the Role of Social Self and threat based concerns. *Journal of Personality and Social Psychology*. 91, 243-254.
- Meidlinger, S. (2003). *Spezifisches Stereotype-Priming und der Einfluss auf die Leistung in einem Fähigkeitstest*. Unveröffentlichte Diplomarbeit Universität Wien.
- Miller, D. T., Taylor, B., & Buck, M. L. (1991). Gender gaps: Who needs to be explained? *Journal of Personality and Social Psychology*, 61, 5–12.
- Mitulla, C. (1997). *Die Barriere im Kopf: Stereotype und Vorurteile bei Kindern gegenüber Ausländern*. Opladen: Leske und Budrich.
- Murphy, S. T. & Zajonc, R. B. (1993). Affekt, cognition, and awareness: Affective Priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64, 723 – 739.
- Nguyen, H.-H. D. & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93, 1314-1334.
- O'Brien, L., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, 29, 782–789.

- OECD (2004). *Learning for Tomorrow's World – First Results from PISA 2003*, OECD, Paris.
 WWW:<http://www.oecd.org/edu/school/programmeforinternationalstudien/tassessmentpisa/33917683.pdf> (abgerufen am 13.06.2015).
- OECD (2013). *PISA 2012 Ergebnisse im Fokus: Was 15-Jährige wissen und wie sie dieses Wissen einsetzen können*, Berlin.
 WWW:<http://www.oecd.org/berlin/themen/PISA-2012-Zusammenfassung.pdf> (abgerufen am 16.06.2015).
- Osborne, J. W. & Walker, C. (2006). Stereotype Threat, Identification with Academics, and Withdrawal from School: Why the most successful students of colour might be most likely to withdraw. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 26, (4), 563-577.
- Ottensschläger, J., Ratzinger, W., Vormayr, D. & Binder, J. (1999). Mathematik 4 – 8. Schulstufe (3. Auflage). Linz: Veritas-Verlag.
- Patterson, M.L. & Werker, J. F. (2002). Infants' Ability to Match Dynamic Phonetic and Gender Information in the Face and Voice. *Journal of Experimental Child Psychology*, 81, 93–115.
- Powlishta, K.K. (1995). Intergroup processes in childhood: Social categorization and sex role development. *Developmental Psychology*, 31, 781-788.
- Quinn, D. M. & Spencer, S. J. (2001). The interference of stereotype threat on woman's generation of mathematical problem solving strategies. *Journal of Social Issues*, 57, 55-71.
- Reichel, H.C., Müller, R., Laub, J., Hanisch, G., Körperth, W., Bernhard, M., Groß, H., Hederer, O. & Wenzel, C. (1999). Lehrbuch der Mathematik, 5. Band (4. Auflage). Wien: öbv & hpt.
- Reischer, R. (2015). *Der Einfluss der Übernahme eines geschlechtsstereotypen Rollenbildes auf die Mathematikleistungen von Jugendlichen*. Unveröffentlichte Diplomarbeit, Universität Wien.

- Robinson C.C. & Morris J.T. (1986). The gender-stereotyped nature of Christmas toys received by 36-, 48-, and 60-month-old children: A comparison between nonrequested vs requested toys. *Sex Roles*, 15, 21-32.
- Scafidi, T. & Bui, K. (2010). Gender similarities in math performance from middle school through high school, *Journal of Instructional Psychology*, 37, (3), 252–255.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 10, (2), 120-148.
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38, 194–201.
- Schmader, T. & Johns, M. (2003). Converging Evidence That Stereotype Threat Reduces Working Memory Capacity. *Journal of Personality and Social Psychology*, 85, (3), 440–452.
- Schmader, T., Johns, M. & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles: A Journal of Research*, 50, 835–850.
- Schmader, T., Johns, M. & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115, 336-356.
- Schneider, D. J. (2004). *The Psychology of Stereotyping*. New York: Guilford.
- Shih, M., Pittinsky, T. L. & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10, 80–83.
- Snyder, M. & Miene, P. (1994). *On the functions of stereotypes and prejudice*. In: *The Psychology of Prejudice*. The Ontario symposium, Zanna MP, Olson JM (eds). Erlbaum: Hillsdale, NJ.

- 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.
- Städtler, T. (2003). *Lexikon der Psychologie*. Stuttgart: Kröner.
- Stangl, W. (2011). *Geschlechtstypisierung*. Lexikon für Psychologie und Pädagogik. WWW: <http://lexikon.stangl.eu/3568/geschlechtstypisierung/> (abgerufen am 25.08.2015).
- Stangl, W. (2011). *Trait*. Lexikon für Psychologie und Pädagogik. WWW: <http://lexikon.stangl.eu/4752/trait/> (abgerufen am 25.08.2015).
- Steele, C. M. (1997). A threat in the air: How Stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629.
- Steele, C.M., & Aronson, J. (1995). Stereotype Threat and the intellectual test-performance of African-Americans. *Journal of personality and Social Psychology*, 69, (5), 797-811.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology*. 34, 379 – 440.
- Stone, J., Lynch, C. I., Sjomeling, M. & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, 77, 1213-1227.
- Tomasetto, C., Alparone, F. R. & Cadinu, M. (2011). Girls' math performance under stereotype threat: the moderating role of mothers' gender stereotypes. *Developmental Psychology*, 47, (4), 943-949.
- Trautner, H.M. (1991). *Lehrbuch der Entwicklungspsychologie* Bd. 2: Theorien und Befunde. Göttingen: Hogrefe.
- Walton, G., & Cohen, G. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39, (5), 456-467.

- Wendt, K. (2015). *Expect the Expected. Stereotype – so entstehen sie, selten gehen sie: Deutsche Stereotype in den USA*. Hamburg, Diplomica Verlag GmbH 2015.
- Weis, R. H. (2006). *Grundintelligenztest Skala 2 - Revision - (CFT 20-R)*. Gottingen: Hogrefe.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245–271.

Tabellenverzeichnis

- Tabelle 1: Aufteilung der Studienteilnehmer auf die verschiedenen Versuchsbedingungen. Hypothesenprüfung.....S.59
- Tabelle 2: Reliabilitätstabelle für die vorgegebenen CFT20 R Matrizen Aufgaben 6 bis 15. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben.....S.60
- Tabelle 3: Aufgabenschwierigkeiten der CFT 20 R Matrizen Aufgaben. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben.....S.61
- Tabelle 4: Deskriptive Statistiken des Pre-Priming Mathematiktest.....S.62
- Tabelle 5: Aufgabenschwierigkeiten der acht Aufgaben des Pre-Priming Mathematiktests.....S.63
- Tabelle 6: Reliabilitätstabelle des Pre-Priming Mathematiktests, welcher nach dem Priming vorgegeben wurde. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben.....S.63
- Tabelle 7: Deskriptive Statistiken des Mathematiktest nach dem Priming.....S.64
- Tabelle 8: Reliabilitätstabelle des Post-Priming Mathematiktest wurde. Pro Aufgabe wird genau ein Punkt für eine richtige Lösung vergeben.....S.65
- Tabelle 9: Aufgabenschwierigkeit der acht Aufgaben des Mathematiktests nach dem Priming.....S.65
- Tabelle 10: Deskriptive Statistiken und Ergebnisse des t-Test für unabhängige Stichproben und Levene's Test der weiblichen und männlichen Teilnehmer bezüglich des Pre-Priming Mathematiktest.....S.69
- Tabelle 11: Angegebene Mathematiknoten der männlichen und weiblichen Teilnehmer im Vergleich.....S.69

- Tabelle 12: Rangdaten der MathematiknotenS.69
- Tabelle 13: Ergebnisse des t-Test und des Levene's Test für die erreichten Punkte im CFT Untertest MatrizenS.70
- Tabelle 14: Varianzanalyse mit Messwiederholung und Levene's Test der Mathematiktests unter Berücksichtigung des Primings und des Geschlechts aller Teilnehmer mit vollständigen Datensets.....S.71
- Tabelle 15: Mittelwerte und Standardabweichungen der erreichten Punkte im Mathematiktest vor und nach dem Priming aufgeteilt auf ihre PrimingbedingungenS.72
- Tabelle 16: Varianzanalyse mit Messwiederholung und Levene's Test unter Ausschluss der Personen, welche die Frage nach dem Inhalt des Artikels falsch beantwortetenS.73
- Tabelle 17: Deskriptive Statistiken der erreichten Punkte im Pre-Priming Mathematiktest und Post-Priming Mathematiktest unter Ausschluss der Personen, welche die Frage nach dem Inhalt des Artikels falsch beantwortetenS.73
- Tabelle 18: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.74
- Tabelle 19: Glaubwürdigkeitseinschätzung des Artikels seitens der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.75
- Tabelle 20: Mediane und Quartile der Arbeitshaltungen der weiblichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.76
- Tabelle 21: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.76

- Tabelle 22: Glaubwürdigkeitseinschätzung des Artikels seitens der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.77
- Tabelle 23: Mediane und Quartile der Arbeitshaltungen der weiblichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.78
- Tabelle 24: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming, sowie das Ergebnis des t-Test der männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.79
- Tabelle 25: Glaubwürdigkeitseinschätzung des Artikels seitens der männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.79
- Tabelle 26: Mediane und Quartile der Arbeitshaltungen männlichen Teilnehmer, welche das stereotypkonforme Priming durchliefenS.80
- Tabelle 27: Deskriptive Glaubwürdigkeitseinschätzung des Artikels seitens der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.81
- Tabelle 28: Deskriptive Statistiken der erreichten Punkte in den Mathematiktests vor und nach dem Priming sowie das Ergebnis des t-Test der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.82
- Tabelle 29: Median und Quartile der männlichen Teilnehmer, welche das stereotypnonkonforme Priming durchliefenS.83
- Tabelle 30: Ergebnisse des t-Tests und Levene's Test von männlichen und weiblichen Teilnehmern, welche das ST Priming erhalten hattenS.84

- Tabelle 31: Ergebnis der einseitigen Spearman Korrelation zwischen der persönlichen Wichtigkeit und der Einschätzung der erreichten Punkte im Baselinetest MathematiktestS.85
- Tabelle 32: Deskriptive Statistiken der eigenen Einschätzungen bezüglich der erreichten Punktezahl in dem Mathematiktest vor und dem Mathematiktest nach dem PrimingS.87
- Tabelle 33: Ergebnisse der Varianzanalyse mit Messwiederholung und Levene's Test bezüglich der Selbsteinschätzung der erreichten Punkte in den Mathematiktests vor bzw. nach dem PrimingS.88
- Tabelle 34: Varianzanalyse mit Messwiederholung und Levene's Test der Fremdeinschätzungen der erreichten Punkte in den Mathematiktests von Burschen vor und nach dem PrimingS.90
- Tabelle 35: Varianzanalyse mit Messwiederholung und Levene's Test der Fremdeinschätzungen der erreichten Punkte in den Mathematiktests von Mädchen vor und nach dem PrimingS.91
- Tabelle 36: Mittelwerte, Standardabweichungen und Anzahl der Einschätzungen von Mädchen richtig gelösten Aufgaben in den Mathematiktests vor und nach dem PrimingS.92
- Tabelle 37: Median und Anzahl der Einschätzungen weiblichen und männlichen Teilnehmer bezüglich mathematischer Fähigkeiten von dem Pre-Priming Mathematiktest und dem Post-Priming MathematiktestS.95
- Tabelle 38: Median und Quartile der Arbeitshaltungen bei dem Termin an dem die Versuchspersonen mittels Artikel geprimt wurdenS.96
- Tabelle 39: Ergebnisse des U-Tests bezüglich der ArbeitshaltungenS.97
- Tabelle 40: Faktorenanalyse: Fragebogen ArbeitshaltungenS.144
- Tabelle 41: Anzahl der Teilnehmer unter Ausschluss jener Teilnehmer mit einer geringen Motivation.....S.144

- Tabelle 42: Varianzanalyse des Haupteffekts unter Ausschluss der Teilnehmer mit einer geringen Motivation.....S.144
- Tabelle 43: Levene's Test für die Mathematikleistungen des Pre- und Post Priming Mathematiktests unter Ausschluss der Teilnehmer mit einer geringen MotivationS.145
- Tabelle 44: Ergebnisse der Levene's Tests für die männlichen Teilnehmer, der ST und SL Bedingungen bezüglich der erreichten Punkte im Pre- Priming und Post-Priming Mathematiktest.....S.145
- Tabelle 45: Levene's Test für den Post-Priming Mathematiktest weibliche Teilnehmer in der ST und SL Bedingung.....S.145
- Tabelle 46: Deskriptive Statistiken der Einschätzung der, von Burschen richtig beantworteter ItemsS.146
- Tabelle 47: Ergebnisse einer Univariaten Varianzanalyse für die Einschätzung der von Mädchen richtig beantworteten Aufgaben im Mathematiktest nach dem PrimngS.146

Abbildungsverzeichnis

- **Abbildung 1:** Integriertes Prozessmodell der Wirkung von Stereotype Threat auf Leistung von Schmader, Forbs und Johns, (2008, S.337)
.....S. 30
- **Abbildung 2:** Ein Screenshot der ersten Beispielaufgabe des CFT 20 R Matrizen Tests mit dazugehöriger Instruktion.....S.46
- **Abbildung 3:** Ein Screenshot des Fragebogens zum Arbeitsstil.....S.50
- **Abbildung 4:** Eine schematisierte Darstellung des ersten und zweiten Teils des Untersuchungsablaufes.....S.52
- **Abbildung 5:** Mögliche Kombinationen der Primingversionen und des Geschlechts der Testperson über die beiden Termine verteilt.....S.55
- **Abbildung 6:** Screenshot des Startbildschirms für die Eingabe des persönlichen Codes.....S.58
- **Abbildung 7:** Ein Histogramm der Verteilung der von den Teilnehmern im Baseline Mathematiktest erreichten Punkte.....S.62
- **Abbildung 8:** Ein Histogramm der Verteilung der von den Teilnehmern erreichten Punkte.....S.64
- **Abbildung 9:** Histogramm, welches die Motivation der weiblichen Teilnehmer mit stereotypnonkonformen Priming zeigt.....S.75
- **Abbildung 10:** Histogramm, welches die Motivation der weiblichen Teilnehmer mit stereotypkonformem Priming zeigt.....S.77
- **Abbildung 11:** Histogramm, welches die Motivation der männlichen Teilnehmer mit stereotypkonformen Priming zeigt.....S.80
- **Abbildung 12:** Histogramm welches die Motivation der männlichen Teilnehmer mit stereotypnonkonformen Priming zeigt.....S.82
- **Abbildung 13:** Interaktionsdiagramm der Selbsteinschätzungen erreichter Punkte im Mathematiktest nach dem Priming für männliche und weibliche Teilnehmer in den verschiedenen Primingbedingungen.....S.89
- **Abbildung 14:** Interaktionsdiagramm der Einschätzung durchschnittlich richtig gelöster Aufgaben im Mathematiktest nach dem Priming für männliche und weibliche Teilnehmer in den verschiedenen Primingbedingungen.....S.93
- **Abbildung 15:** Histogramm der eigenen Einschätzung der Motivation der Teilnehmer.....S.94
- **Abbildung 16:** Angegebene Schwierigkeit aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.140

- **Abbildung 17:** Angegebene Konzentration aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.140
- **Abbildung 18:** Angegebene Geschwindigkeit aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.141
- **Abbildung 19:** Angegebene Forderung aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.141
- **Abbildung 20:** Angegebene Bemühung aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.142
- **Abbildung 21:** Angegebene Interesse aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest.....S.142

Anhang

Anhang A – Abstract

This study was intended to compare effects of different stereotype priming methods on the mathematical performances of high school students. Two manipulation methods were used. The first was an article stating that either males or female do better in mathematical fields, the second a questionnaire consisting of adjective pairs typically associated with females or males. Students were asked to answer the questionnaire from the perspective of a typical male or female. This diploma thesis concentrates on the results of the first priming method. The mathematical ability of students prior to the manipulation and after a manipulation was measured. Mathematic scores before and after priming and between different priming conditions were compared. Results show no gender differences in mathematic scores prior to priming, nor after priming. All participants scored less points in the mathematics test after priming, regardless of the priming condition they had been assigned. Priming conditions had no significant impact on the achieved points in Mathematics test, or the self-assessments.

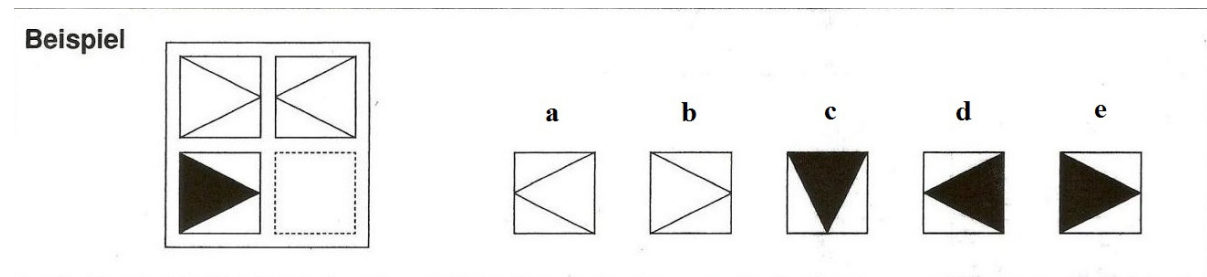
Anhang B – Zusammenfassung

Diese Studie versuchte ursprünglich, unterschiedliche Wirkungen von Stereotypenpriming Methoden auf die mathematischen Leistungen von Schülern zu untersuchen. Zwei Manipulationsmethoden wurden verwendet. Einerseits wurde eine direkte Information mittels eines Artikels, welcher über bessere Mathematikleistungen von entweder Frauen oder Männern berichtete, vorgegeben. Die zweite Primingmethode verlangte, das Ausfüllen eines Fragebogens mit typisch weiblichen und männlichen Charaktereigenschaften. Der Fragebogen sollte aus der Sicht eines typischen Burschen oder Mädchens des Alters der Versuchspersonen ausgefüllt werden. Die Erhebung wurde über den Verlauf von zwei Testzeitpunkten durchgeführt, um die Leistungen der Teilnehmer vor und nach dem jeweiligen Priming zu vergleichen. Aus den Ergebnissen war kein signifikanter Geschlechtsunterschied bezüglich der mathematischen Fähigkeiten vor dem Priming, noch nach dem Priming ersichtlich. Alle Teilnehmer erzielten weniger Punkte bei dem Mathematiktest nach dem Priming, als vor dem Priming, wobei diese Leistungsminderung nur bei den männlichen Probanden signifikant war. Es ergab sich kein nachweisbarer Effekt des Primings, weder auf die mathematischen Leistungen der einzelnen Versuchsgruppen, noch auf die Selbsteinschätzungen der Teilnehmer.

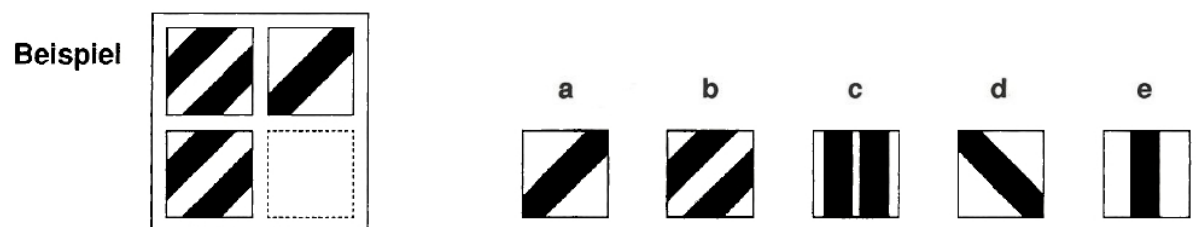
Anhang C – Untersuchungsinstrumente

CFT Aufgaben

- Beispiel 1:

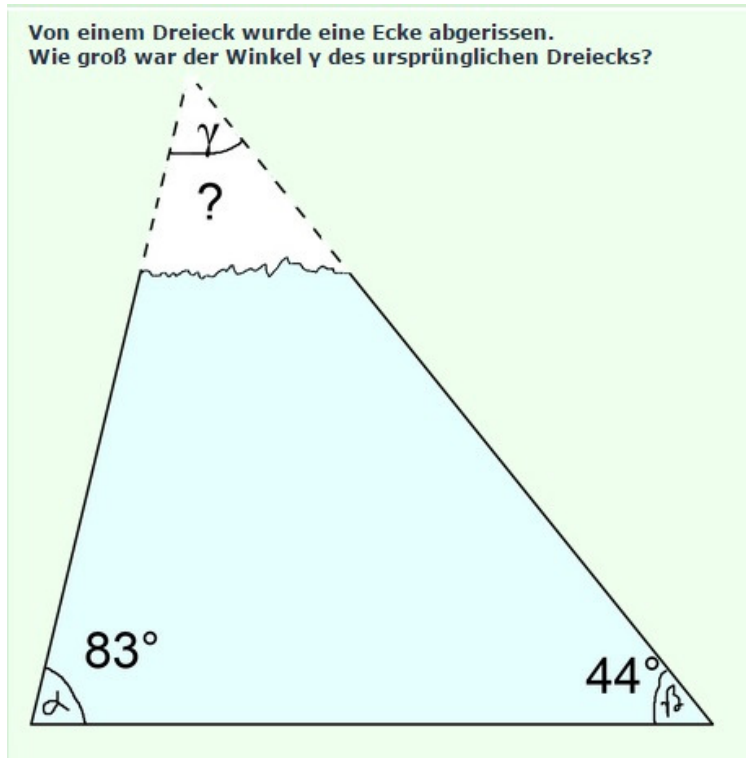


- Beispiel 2:



Mathematiktest Baseline Aufgaben

- Item 1:



- Item 2:

Herr Reisinger kaufte einen Fernseher um 260€. Frau Friedrich zahlte im Ausverkauf für das gleiche Gerät nur 85% des Originalpreises. Wie viel bezahlte sie?

- Item 3:

Frieda und Rainer haben zusammen 60 Stück Nüsse. Wenn jeder 10 Nüsse verschenkt, dann hat Frieda dreimal so viele Nüsse wie Rainer. Wie viele Nüsse hat Frieda?

- Item 4:

Zum Aufbau einer Ausstellung werden 200 Personen gebraucht, wenn die Arbeit in 12 Stunden bewältigt werden soll. Wie viele Personen braucht man, hätte man zum Aufbau 48 Stunden zur Verfügung?

- Item 5:

Frau Moser kauft 15 kg Erdbeeren und 20 kg Marillen zum Einkochen und bezahlt dafür insgesamt 81€. Frau Strasser kauft jeweils um 5 kg weniger und bezahlt daher um 24€ weniger als Frau Moser. Wie viel kostet ein Kilo Marillen?

- Item 6:

Die Summe der Quadrate zweier aufeinander folgender natürlicher Zahlen ist um 243 größer als das Doppelte der kleineren Zahl. Berechne die kleinere Zahl!

- Item 7:

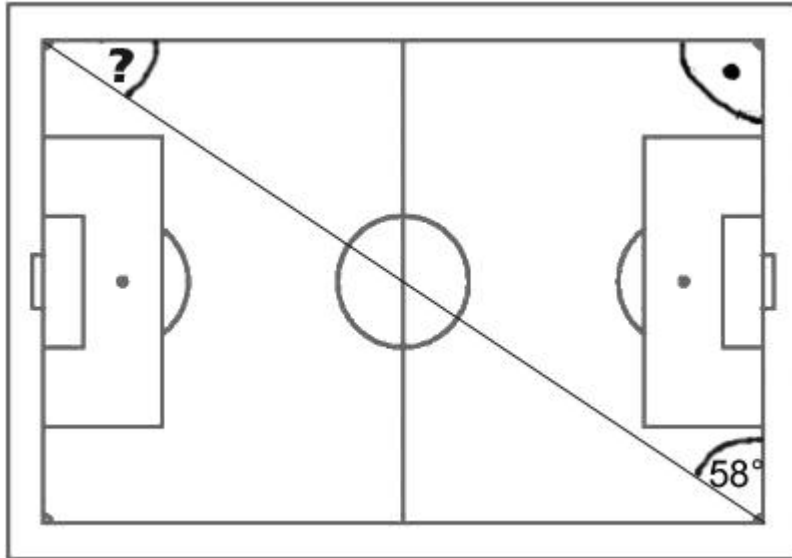
Aus 10 Liter reinem Alkohol (96%) soll durch den Zusatz von Wasser 60%iger Alkohol hergestellt werden. Wie viel Liter Wasser müssen zugesetzt werden?

- Item 8:

Herr Bauer wohnt in Wien, seine Schwester im 225 km entfernten Graz. Er ist zum Mittagessen bei seiner Schwester eingeladen. Wann muss er wegfahren, wenn er um 11:45 in Graz sein möchte und mit einer mittleren Geschwindigkeit von 90km/h unterwegs ist?

Mathematiktest Aufgaben nach dem Priming

- Item 1:



Auf einem Fußballfeld wird eine Linie genau diagonal von einer Ecke zur anderen gezogen. Welche Größe hat der mit "?" gekennzeichnete Winkel wenn das Fußballfeld rechteckig ist?

- Item 2:

Frau Jäger hat bei einem Preisausschreiben 860€ gewonnen. Von diesem Geld legt sie 15% beiseite. Wie viel Euro verbleiben ihr noch?

- Item 3:

Die beiden Superstars der 4a Klasse, Andreas und Christina, haben zusammen 25 „Sehr gut“ bei den Schularbeiten erreicht. Hätte Andreas um 2 „Sehr gut“ mehr, Christina um 1 weniger, so wären beide gleich erfolgreich gewesen. Wie viele „Sehr gut“ hatte Andreas?

- Item 4:

Der Lebensmittelvorrat einer Expedition reicht für 12 Personen 8 Tage. Wie viele Tage würde der Vorrat für 6 Personen reichen?

- Item 5:

Maria und Herwig besuchen eine Konditorei und erhalten folgende Rechnung:

Maria: 1 Kaffee und 2 Stück Kuchen 5 € Herwig: 2 Kaffee und 1 Stück Kuchen 4,60 €. Wie viel kostet ein Stück Kuchen?

- Item 6:

Subtrahiert man vom Produkt zweier Zahlen, die sich um 2 unterscheiden, das Doppelte der kleineren Zahl, so erhält man 529. Berechne die kleinere Zahl!

- Item 7:

Wie viel Liter Wasser muss zu einer 15%igen Salzlösung hinzugefügt werden, damit sie 10%ig wird, wenn 300 Liter Salzlösung vorhanden sind?

- Item 8:

Herr Winter hört um 20:15 Uhr im Verkehrsfunk, dass die Bundesstraße zwölf bei einer Baustelle ab 21:00 Uhr gesperrt werden muss. Er ist noch 66 km von der Baustelle entfernt. Mit welcher Mittleren Geschwindigkeit muss er mindestens fahren, um noch zeitgerecht die Baustelle passieren zu können?

Artikel in Stereotypkonformer Version

Neue Untersuchung zeigt: Es gibt Geschlechtsunterschiede in
mathematischnaturwissenschaftlichen Fähigkeiten

Von Erin A. Goodey

Die heute veröffentlichte Studie einer internationalen Forschergruppe wirft neues Licht auf die Unterschiede zwischen Männern und Frauen. Die Forschergruppe kommt zu dem Schluss, dass Geschlechterunterschiede im Bereich mathematischer Fähigkeiten tatsächlich existieren. Die Ergebnisse zeigen, dass Männer in der Lage sind deutlich bessere Ergebnisse zu erzielen als Frauen. Die aktuelle Studie ist die größte und umfassendste jemals zum Thema publizierte Untersuchung. Die Teilnehmer und Teilnehmerinnen an der Untersuchung wurden über acht Jahre hinweg begleitet und ihre Leistungen wurden genau dokumentiert. An der Studie haben über 50 Millionen Personen teilgenommen, dies ist der größte jemals verzeichnete Stichprobenumfang. Dabei verglichen die Forscher die Leistungen in Mathematik von Männern und Frauen in 113 Ländern rund um die Welt. Die Leistungen wurden mit standardisierten Tests für Grund- und weiterführende Schulen erhoben. Die Ergebnisse zeigen, dass Männer im Schnitt bei Mathematikaufgaben bessere Leistungen erzielen als Frauen. Entsprechende Ergebnisse wurden auch für Studierende an Universitäten in 26 Ländern gefunden. Die Forschung wurde vom National Institute of Health (NIH) gefördert, die das internationale Forscherteam unter der Leitung von Dr. Mark Goldstein vom Harvard Gender Research Institute mit 35 Millionen Dollar unterstützte. Die Ergebnisse wurden heute in „Child Development“, einer führenden Zeitschrift der American Psychology Association, veröffentlicht. Es ist zu erwarten, dass diese produktive Forschergruppe noch viele weitere Artikel publizieren wird. Dr. Thomas Schmidt meinte im Namen des Teams: „Allein der Umfang der Studie garantiert die Richtigkeit der Ergebnisse.“ „Der Umfang der Studie wirft neues Licht auf die alte Diskussion um Geschlechterrollen und mathematische Leistungen. Ich hoffe, dass auch Lehrer und Lernende nun ihre Erwartungen ändern werden“, sagt die Autorin Dr. Karen Dinear, Direktorin des Kinder- und Jugendkrankenhauses an der University of Wisconsin.

Fragen nach dem Artikel (Erinnerungstest)

1.

Die Hauptaussage des Artikels ist...

	Richtig	Falsch
Frauen erzielen bessere Ergebnisse im Mathematischen Bereich als Männer	☐	☐
Männer erzielen bessere Ergebnisse im Mathematischen Bereich als Frauen	☐	☐
Männer und Frauen erzielen gleich gute Ergebnisse im Mathematischen Bereich	☐	☐

2.

***Der Autor des Artikels heißt Erin A. Goodey**
Bitte wähle eine der folgenden Antworten:

- ☐ Richtig
- ☐ Falsch

3.

***An der Studie nahmen 1000 Personen teil.**
Bitte wähle eine der folgenden Antworten:

- ☐ Richtig
- ☐ Falsch

4.

*** Der Artikel erschien in der Zeitschrift „Child Development“
Bitte wähle eine der folgenden Antworten:**

- ☐ Richtig
- ☐ Falsch

5.

***Die Personen welche an der Studie teilnahmen würden über 12 Jahre lang
begleitet.
Bitte wähle eine der folgenden Antworten:**

- ☐ Richtig
- ☐ Falsch

6.

***Das National Institute of Health (NIH) unterstützte das internationale
Forscherteam mit 35 Millionen Dollar.
Bitte wähle eine der folgenden Antworten:**

- ☐ Richtig
- ☐ Falsch

7.

***Glaubst du dass die Ergebnisse des Artikels stimmen?**

- ☐ Ja
- ☐ Nein

8.

***Wie sehr stimmst du mit der Kernaussage des Artikels überein?
Bitte wähle eine der folgenden Antworten:**

- ☐ Gar nicht
- ☐
- ☐
- ☐
- ☐
- ☐ Sehr

Screenshot aus LimeSurvey (Demographische Daten)

*** Welche Sprache sprichst du hauptsächlich Zuhause?**
Bitte wähle eine der folgenden Antworten:
Bitte beantworte diese Frage.
Falls Du "Sonstige:" Angabe möglich:" auswählen solltest, spezifiziere bitte Deine Auswahl im entsprechenden Textfeld.

Bitte auswählen.. ▾

*** Welche Schule hast du im letzten Schuljahr besucht?**
Bitte wähle eine der folgenden Antworten:
Bitte beantworte diese Frage.

☐ Hauptschule
☐ Mittelschule
☐ Gymnasium
☐ Realgymnasium

*** Wie sehr interessierst du dich für**
Bitte beantworte diese Frage. Bitte beantworte alle Bereiche innerhalb der Fragen..

	Gar nicht	.	.	.	.	Sehr
Englisch	☐	☐	☐	☐	☐	☐
Deutsch	☐	☐	☐	☐	☐	☐
Mathematik	☐	☐	☐	☐	☐	☐

Vielen Dank, dass du bei unserer Studie mitmachst. Zu Beginn benötigen wir einige Informationen über dich. Sämtliche Daten, die du eingibst werden anonym behandelt und nur zum Zweck der statistischen Auswertung verwendet. Bitte versuche sorgfältig und zügig zu arbeiten und sämtliche Fragen ehrlich zu beantworten.

*** Geschlecht**
Bitte beantworte diese Frage.

☐ Weiblich ☐ Männlich

*** Geburtsmonat**
Bitte beantworte diese Frage.
Jede Antwort muss zwischen 1 und 12 sein

In diesem Feld sind nur numerische Werte erlaubt

*** Geburtsjahr**
Bitte beantworte diese Frage.

In diesem Feld sind nur numerische Werte erlaubt

*** Welche Note hattest du im letzten Zeugnis in**

**Bitte beantworte diese Frage. Bitte beantworte alle Bereiche innerhalb der Fragen..
Jede Antwort muss zwischen 1 und 6 sein**

In diese Felder dürfen nur Ziffern eingetragen werden

Englisch	
Deutsch	
Mathematik	

Zusatzblatt Matheaufgaben

Mein persönlicher CODE:

Erster Buchstabe deines Vornamens	Dritter Buchstabe des Mädchennamens deiner Mutter (Nachname vor der Hochzeit)	Letzte Ziffer der Hausnummer in der du wohnst	Dritter Buchstabe deines Geburtsmonats
(z.B. Maria → M)	(z.B. Mustermann → S)	(z.B. Mustergasse 129 → 9)	(z.B. November → V)
—	—	—	—

1.

2.

3.

4.

1.

2.

3.

4.

5.

6.

7.

8.

Anhang D – Histogramme

Histogramme der Arbeitshaltungen

Abbildung 16: Angegebene Schwierigkeit aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest

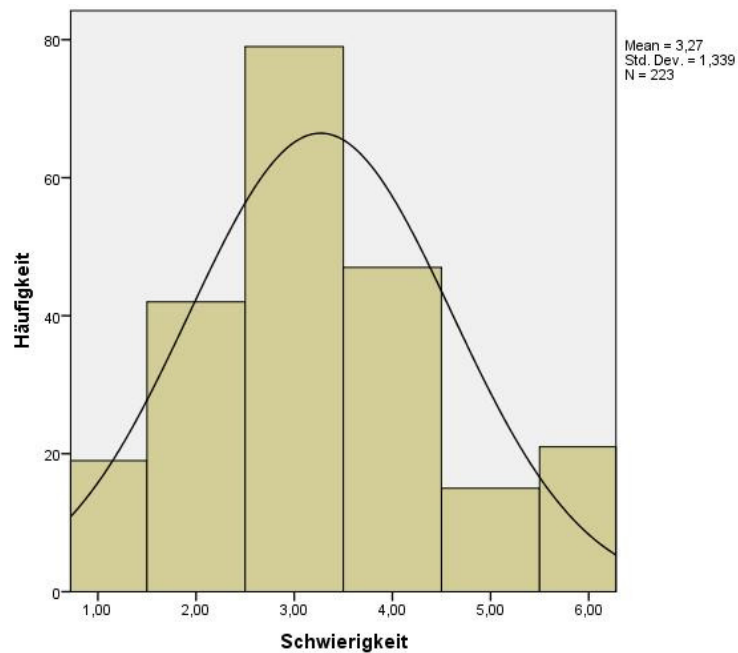


Abbildung 17: Angegebene Konzentration aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest

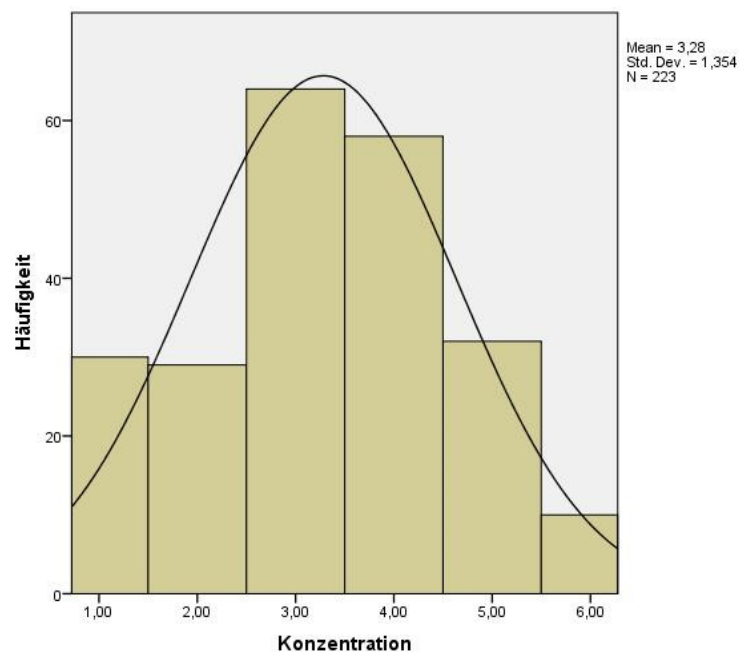


Abbildung 18: Angegebene Geschwindigkeit aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest

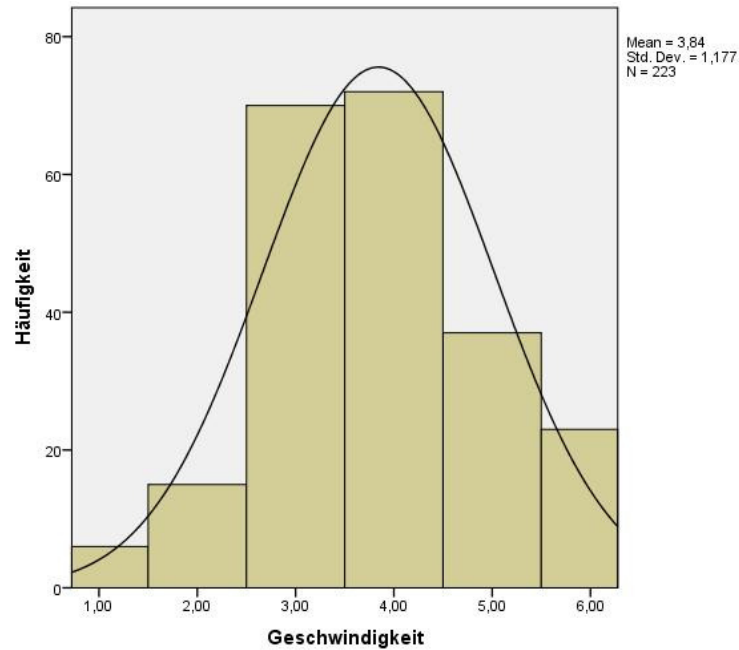


Abbildung 19: Angegebene Forderung aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest

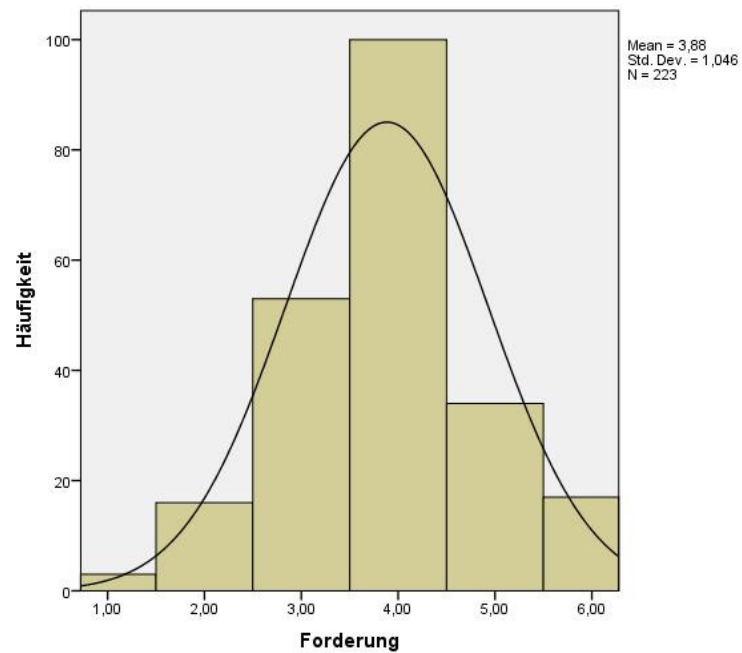


Abbildung 20: Angegebene Bemühung aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest

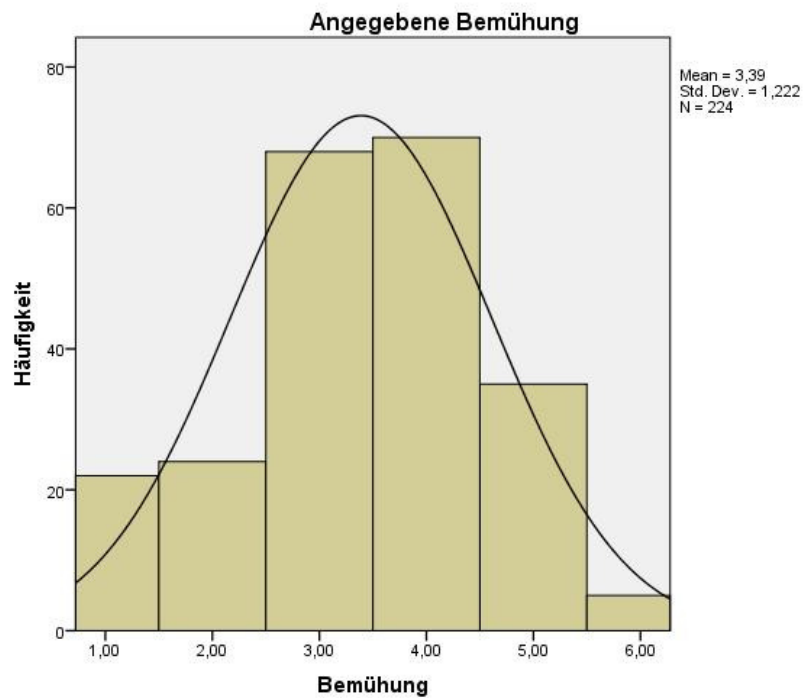
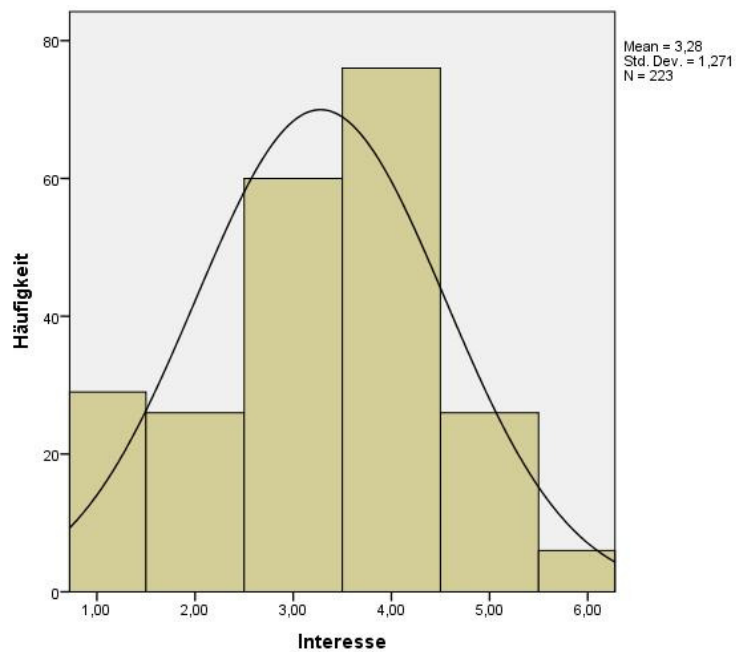


Abbildung 21: Angegebene Interesse aller bei beiden Terminen anwesenden Teilnehmer bei dem Post- Priming Mathematiktest



Anhang E – Tabellen

Tabelle 40: Faktorenanalyse: Fragebogen Arbeitshaltungen

	Testzeitpunkt			
	1		2	
	Komponente		Komponente	
	1	2	1	2
unkonzentriert vs. konzentriert	0.804	0.091	0.800	-0.078
langsam vs. schnell	-0.012	0.698	-0.002	0.592
nicht interessiert vs. interessiert	0.790	0.069	0.840	0.065
unmotiviert vs. motiviert	0.794	0.086	0.805	0.198
überfordert vs. unterfordert	0.046	0.803	0.070	0.785
wenig bemüht vs. sehr bemüht	0.843	-0.162	0.860	-0.123
Rechenaufgaben waren schwierig vs. einfach	0.057	0.794	-0.040	0.862

Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung.
 a Die Rotation ist in 3 Iterationen konvergiert.

Varianzanalyse des Haupteffekts unter Ausschluss der Teilnehmer mit einer geringen Motivation

Tabelle 41: Anzahl der Teilnehmer unter Ausschluss jener Teilnehmer mit einer geringen Motivation

		N
Geschlecht	Weiblich	76
	Männlich	70
Priming	Artikel „Frauen schneiden besser in Mathematischen Bereichen ab“	77
	Artikel „Männer schneiden besser in Mathematischen Bereichen ab“	69

Tabelle 42: Varianzanalyse des Haupteffekts unter Ausschluss der Teilnehmer mit einer geringen Motivation

	F	Sig.	MQ	df
Testleistung	6.006	0.015	6.824	1
Testleistung * Geschlecht	0.417	0.520	0.473	1
Testleistung * Priming	0.119	0.731	0.135	1
Testleistung * Geschlecht * Priming	0.601	0.439	0.683	1
Geschlecht	0.745	0.390	2.625	1
Priming	0.192	0.662	0.676	1
Geschlecht * Priming	0.097	0.756	0.343	1

Tabelle 43: Levene's Test für die Mathematikleistungen des Pre- und Post Priming Mathematiktests unter Ausschluss der Teilnehmer mit einer geringen Motivation

Levene Test	F	Sig	df1	df2
Pre-Priming Mathematiktest	1.814	0.147	3	142
Post – Priming Mathematiktest	0.596	0.619	3	142

Tabelle 44: Ergebnisse der Levene's Tests für die männlichen Teilnehmer, der ST und SL Bedingungen bezüglich der erreichten Punkte im Pre-Priming und Post-Priming Mathematiktest

Levene Test	F	Sig	df1
Pre-Priming Mathematiktest	0.581	0.004	104
Post – Priming Mathematiktest	8.538	0.057	98.494

Tabelle 45: Levene's Test für den Post-Priming Mathematiktest weibliche Teilnehmer in der ST und SL Bedingung

Levene Test	F	Sig	df
Post – Priming Mathematiktest	0.843	0.361	117

Tabelle 46: Deskriptive Statistiken der Einschätzung der, von Burschen richtig beantworteter Items

	Geschlecht	Priming	Mittelwert	St. Abweichung	N
Fremdeinschätzung Burschen Pre-Priming Mathematiktest	Weiblich	Stereotypnonkonform	4.721	1.451	61
		Stereotypkonform	5.431	1.365	58
		Gesamt	5.067	1.448	119
	Männliche	Stereotypnonkonform	4.746	1.590	55
		Stereotypkonform	4.804	1.613	51
		Gesamt	4.773	1.593	106
	Gesamt	Stereotypnonkonform	4.733	1.511	116
		Stereotypkonform	5.138	1.512	109
		Gesamt	4.929	1.522	225
Fremdeinschätzung Burschen Post-Priming Mathematiktest	Weiblich	Stereotypnonkonform	4.51	1.523	61
		Stereotypkonform	5.12	1.612	58
		Gesamt	4.81	1.591	119
	Männlich	Stereotypnonkonform	4.35	2.319	55
		Stereotypkonform	4.49	2.139	51
		Gesamt	4.42	2.225	106
	Gesamt	Stereotypnonkonform	4.43	1.935	116
		Stereotypkonform	4.83	1.895	109
		Gesamt	4.62	1.921	225

Tabelle 47: Ergebnisse einer Univariaten Varianzanalyse für die Einschätzung der von Mädchen richtig beantworteten Aufgaben im Mathematiktest nach dem Priming

	MQ	F	Sig.	df
Geschlecht	31.558	9.352	0.002	1
Total				225

Anhang F – Lebenslauf

Persönliche Daten

Name Janine Martens
E-Mail janine.martens@hotmail.com
Geburtsdatum 12.12.1983 in Hamburg
Staatsangehörigkeit Deutsch



Schulbildung

1986 - 1988 **Kindergarten** an der Internationalen Schule Hamburg
1988 - 2002 **Preschool bis Senior School** an der Internationalen Schule Hamburg
05/2002 Abschluss des International Baccalaureate's und des High School Diploms an der Internationalen Schule Hamburg
10/2002 – 03/2003 Graphiker Ausbildung an der Kunstschule Wandsbek, Hamburg
10/2003 – 01/2016 Studium Psychologie (Diplom) an der Universität Wien
01/2009 Vordiplom Psychologie an der Universität Wien

Berufliche Laufbahn

10/2002 – 01/2003 Qualitätssicherung (Teilzeit)
Troge Medical GmbH, Hamburg
07/2005 – 09/2005 Datenauswertung
ESSO Deutschland GmbH, Hamburg
07/2010 – 09/2010 Universitäres Praktikum
Klinik für Forensische Psychiatrie und Psychotherapie,
Lüneburg
03/2012 – 12/2012 Praktikum und Assistenz Tätigkeiten
Psychologische Praxis Dr. Evelyn Patzak, Wien
Assistenz bei psychologischen Testungen und
Gutachtenerstellung

Sprachkenntnisse:

Muttersprache Deutsch
Fließend in Wort und Schrift Deutsch, Englisch