

**Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin**

„Schlaue Mädchen – dumme Jungs?!“
Der Einfluss von negativen Stereotypen über Jungen auf
Leistung, Motivation und Gruppenprozesse im Klassenzimmer.

**Dissertation zur Erlangung des akademischen Grades
Doktor der Philosophie (Dr. phil.)**

vorgelegt von:
Martin Latsch, M.Sc.

Berlin, 2014

Erstgutachterin: Prof. Dr. Bettina Hannover

Zweitgutachterin: Prof. Dr. Felicitas Thiel

Tag der mündlichen Prüfung: 01. Oktober 2014

Diese Arbeit entstand auf Grundlage des von der Deutschen Forschungsgemeinschaft geförderten Projekts „Sind Jungen die neuen Bildungsverlierer?“ unter Antragstellung von Prof. Bettina Hannover (HA-2381-11/1&2) und Prof. Ursula Kessels (KE-1412-2/1&2).

DANKSAGUNG

An dieser Stelle sei all den Personen von ganzem Herzen gedankt, welche immer wieder Mut und Zuspruch gegeben und fachlich zur Seite gestanden haben.

Zuerst sei den Gutachtenden für Ihre Mühe gedankt, besonders Frau Prof. Dr. Bettina Hannover für die zahllosen fachlichen Diskussionen, die angenehme Arbeitsatmosphäre, das Teilen jeder Probleme und das häufige Zuhören. Ohne das wäre es mir unmöglich gewesen, diese Arbeit zu verfassen. Ich weiß sehr zu schätzen, was du für mich getan hast – hoffentlich gelingt es mir, nur einen Teil davon wiederzugeben. Frau Prof. Dr. Thiel gilt mein herzlichster Dank für die sofortige und kurzfristige Bereitschaft zur Übernahme der Zweitbegutachtung und die dafür aufgewendeten Stunden, ebenso der gesamten Promotionskommission, Prof. Martin Brunner, Prof. Ursula Kessels und Dr. Lysann Zander, wo Zeit doch mittlerweile eine so knappe Ressource geworden ist.

„Meinem“ Team des Arbeitsbereichs Schul- und Unterrichtsforschung und des DFG-Projekts möchte ich ebenfalls für den Rückhalt danken, der mir während des Erstellens dieser Arbeit gegeben wurde. Neben Kollegen/-innen habe ich hier auch Freunde gefunden und bin sehr glücklich, ein Teil von euch zu sein – ohne Franziska Moser, Sara Matthies (*„Ich habe was zu besprechen – lasst uns eine rauchen gehen“* ©), Dörte Adam, Anke Heyder, Franziska Böttcher, Ursula Beermann, Lysann Zander und Janine Neuhaus wäre so mancher Tag verloren gegangen, den wir gemeinsam wieder retten konnten. Ein besonderer Dank gilt auch den studentischen Mitarbeiterinnen und Mitarbeitern, die so manche Erhebung gerettet, manchen Datensatz erstellt und bei so manchem Kaffee mit mir diskutiert haben. Danke Georgina Neuhaus, Michael Schulte, Lisa Gerloff, Klara Reichenbach und Anne Rössner ihr wart (und seid) die „guten Seelen“ des Projekts. Was auch immer uns entgegentrat, ich konnte mich jederzeit auf euch verlassen. So ist diese Arbeit nicht minder eure Arbeit, denn nur durch eure tatkräftige Unterstützung konnte ich mir die Räume erkämpfen, etwas aufs Papier zu bringen.

Im Beruf erfolgreich zu sein hängt allerdings nicht nur von einer hervorragenden Teamleitung und einem exzellenten Kollegium ab, auch in meinem privaten Umfeld schulde ich Dank an so viele Menschen, die an mich geglaubt und gehofft, mich getröstet und motiviert haben. Meinem Partner André Nowakowski für die vielen Stunden, welche die Arbeit mich schier verschlang, das Verständnis, die Liebe und die Zuversicht, wenn ich sie nicht mehr zu haben vermochte. Ohne diese Quelle täglicher Kraft und diese sichere Rückzugsmöglichkeit wäre ich auf dem Weg zum Doktorgrad verloren gegangen, wenn du mich nicht stets wieder gefunden hättest. Sabine Kühn, Elmar Fust, Tero und dem Rest des Rudels für die Zeiten vollkommener Entspannung. Dr. Martyn Vilain fürs „Aufgleisen“ aus verfahrensten Situationen. Meinen Eltern Bärbel und Andreas Latsch, meiner Schwester Susann samt Familie und meinen lieben Großeltern Edelgard (*„Wann bist du endlich fertig?“*) und Helmut Rabold (†) für die schönen Stunden in familiärer Geborgenheit und das Vertrauen, dass ich meinen Weg gefunden habe. Danke euch für Alles!

INHALTSVERZEICHNIS

I. Theoretischer und empirischer Hintergrund	6
1. Einleitung: Sind Jungen die neuen Bildungsverlierer?	6
1.1. Empirische Daten zum Bildungserfolg von Jungen	7
1.2. Entstehung von Geschlechterunterschieden.....	12
1.3. Soziale Konstruktion von Geschlecht in der Schule.....	15
1.4. Erklärungsansätze für Geschlechterdisparitäten im Bildungswesen	21
2. Stereotype Threat-Theorie: Stereotypenbedrohung für Jungen i. d. Schule?	32
2.1. Aktivierung von stereotype threat	35
2.2. Mechanismen von stereotype threat	36
2.3. Verminderung von stereotype threat Effekten	42
2.4. Stereotype reactance.....	44
2.5. Die andere Seite der Medaille: Stereotype lift und stereotype boost.....	49
2.6. Anwendung auf die Problemstellung	50
3. Jenseits von Leistung: Stereotypenbedrohung und Motivation	53
3.1. Theoretische Grundlagen	53
3.2. Empirische Befunde.....	56
3.3. Das Stereotyped Task Engagement Process-Modell	63
3.4. Anwendung auf die Problemstellung	70
4. Das Soziale als Teil des Selbst: Die Theorie der sozialen Identität	72
4.1. Entstehungszusammenhang	72
4.2. Grundannahmen der Social Identity Theory: Theoretische Kontinua.....	75
4.3. Verhaltenskonsequenzen auf Intergruppenebene	77
4.4. Bedeutung und empirische Überprüfung der Social Identity Theory.....	80
4.5. Anwendung auf die Problemstellung	86
II. Empirischer Teil	89
5. Studie 1: Wie werden Jungen wahrgenommen?	89
5.1. Überblick und Fragestellung.....	90
5.2. Methode	92
5.3. Ergebnisse.....	95
5.4. Diskussion.....	98
6. Studie 2: Stereotype Threat und Leistung bei Jungen	105
6.1. Überblick und Fragestellung.....	106
6.2. Methode	107
6.3. Ergebnisse.....	111

6.4.	Diskussion.....	117
7.	Studien 3 & 4: Stereotype Threat und motivationale Zielorientierungen	123
7.1.	Studie 3: Überblick und Fragestellung.....	124
7.2.	Methode	125
7.3.	Ergebnisse.....	129
7.4.	Studie 4: Replikation	132
7.5.	Gesamtdiskussion.....	136
8.	Studie 5: Wiederherstellung positiver Distinktheit I	141
8.1.	Überblick und Fragestellung.....	141
8.2.	Methode	144
8.3.	Ergebnisse.....	147
9.	Studie 6: Ergänzungsstudie zur Abwertung von Weiblichkeit.....	153
9.1.	Methode	153
9.2.	Ergebnisse.....	155
9.3.	Gesamtdiskussion Studien 5 & 6.....	156
10.	Studie 7: Wiederherstellung positiver Distinktheit II.....	164
10.1.	Überblick und Fragestellung.....	164
10.2.	Methode	165
10.3.	Ergebnisse.....	168
10.4.	Diskussion.....	172
11.	Studie 8: Wiederherstellung positiver Distinktheit III	178
11.1.	Überblick und Fragestellung.....	178
11.2.	Methode	180
11.3.	Ergebnisse.....	181
11.4.	Diskussion.....	186
III.	Gesamtdiskussion	192
	Zusammenfassung	208
	Summary	210
	Literaturverzeichnis	212
	Publikationsverzeichnis.....	236
	Erklärung.....	237
	Appendix	238

I. THEORETISCHER UND EMPIRISCHER HINTERGRUND

1. Einleitung: Sind Jungen die neuen Bildungsverlierer?

Durch Ergebnisse der großen Schulleistungsstudien wie PISA, PIRLS und TIMMS rückten seit der Jahrtausendwende vermehrt Jungen in den Fokus der Ungleichheitsforschung in der Schule und leiteten die sogenannte „Jungenwende“ (Weaver-Hightower, 2003) ein. Das Ziel der vorliegenden Arbeit soll sein, zu identifizieren, ob Jungen im heutigen Schulalltag im Gegensatz zu früheren Kohorten, in welchen der Forschungsfokus verstärkt auf ungleiche Chancen von Mädchen besonders im mathematisch-naturwissenschaftlichen Bereich lag, eine neue Gruppe benachteiligter Akteure darstellt. Darüber hinaus soll geprüft werden, wie Jungen mit diesem veränderten Bild, in welchem Jungen ggf. Daten-Dem vorangestellt ist die initiale Fragestellung, ob entsprechende Vorstellungen des Jungen als Verlierer im deutschen Schulsystem bereits internalisiert vorliegen. Im Ausblick wird darauf fokussiert, inwieweit die Positionierung von Jungen im schulischen Kontext verbessert werden kann und welche Voraussetzungen Jungen für eine erfolgreiche Teilhabe am Bildungssystem benötigen.

Zu Beginn wird ein Überblick über statistische Kennzahlen zum Bildungserfolg von Jungen gegeben und ein ergänzender Blick auf die aktuelle, mediale Debatte um Jungen als Bildungsverlierer geworfen. Zusätzlich wird auf Aspekte der Entstehung von Geschlechterunterschieden und Erklärungsansätze für Geschlechterdisparitäten im Bildungssystem eingegangen. Darauf aufbauend wird in das Konzept der Stereotypenbedrohung (z.B. Steele, 1997) eingeführt und eine Übersicht über vorliegende, sozialpsychologische Forschung zu dieser Theorie (Kapitel 2) gegeben sowie auf Motivation als Konstrukt jenseits von (schulischer) Leistung übertragen (Kapitel 3). Als Reaktion auf die Bedrohung durch das Stereotyp des männlichen Schulversagers schlägt die Theorie der sozialen Identität von Tajfel und Turner (1986) mögliche Verhaltenskonsequenzen zur Wiederherstellung einer positiven Distinktheit der Eigengruppe vor, welche im darauffolgenden Abschnitt (Kapitel 4) erläutert werden.

Basierend auf diesen theoretischen Grundlagen wird in zwei ersten Studien überprüft, ob ein Stereotyp des „Jungen als Schulversager“ bereits im Feld vorzufinden ist (*Studie 1*) und ob eine Aktivierung dieses Stereotyps zur Bedrohung und damit einhergehend zu einer Leistungsver schlechterung in einem standardisierten Leistungstest führt (*Studie 2*). Ergänzend werden zwei Studien beschrieben, welche auf Grundlage von Zielorientierungstheorien (Dweck, 1986) Einflüsse der Debatte um männliche Schulversager auf motivationaler Ebene untersuchten (Studienkomplex 3: *Studien 3 und 4*). Zusätzlich wird aufbauend auf der Theorie der sozialen Identität (Tajfel & Turner, 1986) in vier experimentellen Studien (Studienkomplex 4: *Studien 5 – 8*) untersucht, wie Jungen auf die veränderte soziale Umwelt (Mädchen sind im Bildungssystem

erfolgreicher) reagieren. Es wird überprüft, inwieweit ein direkter sozialer Wettbewerb im Sinne von Aufwertungsprozessen der *Ingroup* der Jungen und Abwertung der *Outgroup* der Mädchen (*Studie 5*), eine Heranziehung außerschulischer Vergleichsdimensionen (*Studie 6*) und eine Änderung der sozialen Vergleichsgruppe (*Studie 7*) von Jungen eingesetzt werden, um die wahrgenommene Überlegenheit der Eigengruppe zu bewahren. Abschließend werden alle Studien in einen Gesamtzusammenhang eingeordnet und kritisch diskutiert, auch hinsichtlich der Möglichkeit zur Ableitung praktischer und bildungspolitischer Implikationen.

1.1. Empirische Daten zum Bildungserfolg von Jungen

Das gesteigerte Forschungsinteresse an Jungen in der Schule basiert einerseits auf Ungleichheiten beim Erwerb fachlicher Kompetenzen, als auch auf deutlichen Unterschieden hinsichtlich der Geschlechterverteilung in verschiedenen Schulformen und dem damit einhergehenden Erwerb von Zertifikaten. Die folgende Darstellung basiert im Wesentlichen auf der Argumentation von Hannover & Kessels (2011), ergänzt um Studien, die auf aktuelleren Daten beruhen. Betreffend den Kompetenzerwerb weisen die Ergebnisse der Internationalen Grundschul-Lese-Untersuchung (IGLU, Bos, Tarelli, Bremerich-Voss & Schwippert, 2012a) darauf hin, dass Jungen bereits am Ende der Klassenstufe 4 signifikant deutlich schlechter lesen können (537 Punkte, entspricht Kompetenzstufe III) als Mädchen (545 Punkte, entspricht Kompetenzstufe IV). Auch im internationalen Vergleich (PIRLS, Mullis, Martin, Kennedy & Foy, 2007) lässt sich dieser Befund in nahezu allen Teilnehmerländern kongruent auffinden. Ergebnisse aus PISA 2012 indizieren, dass dieser Befund am Ende der Klassenstufe 9 noch verstärkt aufzufinden ist: Jungen erreichen mit 486 Punkten noch immer ein deutlich geringeres Niveau an Lesekompetenz als Mädchen mit 530 Punkten. Der internationale Vergleich zeigt weiterhin in allen OECD-Staaten signifikante Geschlechterunterschiede zuungunsten von Jungen mit Nachteilen von 22 bis 62 Kompetenzpunkten (Prenzel, Sälzer, Klieme & Köller, 2013). Auch der Anteil schwächster Leser/-innen ist durch einen erhöht männlichen Anteil geprägt. Auf der untersten und den unteren Kompetenzstufen sind Jungen im Vergleich zu Mädchen deutlich überrepräsentiert (untere Stufe Ib: 0.2% Mädchen, 0.8% Jungen; Stufe Ib: 1.5% Mädchen, 5.0% Jungen; Stufe Ia: 7.0% Mädchen, 14.3% Jungen; OECD, 2014). Die Kompetenznachteile von Jungen im Lesen bestehen darüber hinaus auf allen Subskalen und betreffen somit umfassend alle Bereiche der Lesekompetenz. Besonders bemerkenswert sind diese Befunde vor dem Hintergrund, dass die Lesekompetenz eine bedeutsame Schlüsselkompetenz für das Lernen in sämtlichen Schulfächern darstellt (Bos et al., 2007; Kölbl, Tiedemann & Billmann-Mahecha, 2006; Möller & Schiefele, 2007). Ohne entsprechende Fähigkeiten und Fertigkeiten ist es nur erschwert möglich, sich Lernmaterialien für andere Fächer zu erschließen, da z.B. das Verstehen schriftlicher Anweisungen nur defizitär

möglich ist oder aus Lehrbuchtexten nicht die notwendigen Informationen abgeleitet werden können. Der Nachteil von Jungen auf dieser Dimension erscheint somit schwerwiegend.

Nichtsdestotrotz lassen sich hinsichtlich fachlicher Kompetenzen keineswegs nur Nachteile von Jungen feststellen. Entsprechend den gängigen Geschlechterrollenstereotypen, welche Jungen als besser geeignet für den mathematisch-naturwissenschaftlichen Bereich charakterisieren (Nosek, Banaji & Greenwald, 2002; Steffens & Jelenec, 2011; Steffens, Jelenec & Noack, 2010), erreichen Jungen in der PISA-Studie 2012 (Prenzel et al.) einen Kompetenzvorsprung in der Mathematik, welcher mit 520 Punkten zu 507 Punkten bei Mädchen zwar signifikant, jedoch ungleich geringer ausfällt als der Geschlechtervergleich auf der Skala Lesekompetenz. Darüber hinaus weisen Jungen mit 512 Punkten im Vergleich zu Mädchen mit 505 Punkten einen signifikanten Kompetenzvorsprung im Problemlösen auf und sind in den oberen Leistungsstufen stärker vertreten (Stufe 5: 10.0% Jungen vs. 7.7% Mädchen, Stufe 6: 3.1% Jungen vs. 1.8% Mädchen). Die naturwissenschaftlichen Kompetenzen unterscheiden sich abschließend für die deutsche PISA-Stichprobe 15-jähriger Schüler/-innen mit 518 (Mädchen) zu 523 (Jungen) Punkten nicht bedeutsam voneinander, wenngleich im Grundschulalter die TIMSS-Studie einen signifikanten Leistungsvorteil für Jungen (534 Punkte) im Vergleich zu Mädchen (524 Punkte) feststellt (Bos, Wendt, Köller & Selter, 2012b).

Diese ersten Indikatoren einer Benachteiligung von Jungen werden allerdings in weiteren Forschungsarbeiten deutlich differenzierter betrachtet, als es im Rahmen der PISA-Untersuchung vorgenommen wurde. Während die soeben zitierten Kennwerte aus PISA von Schüler/-innen gemittelt über alle Schulformen gebildet wurden, weisen weitere Analysen aus früheren PISA-Studien unter Kontrolle der Schulform darauf hin, dass innerhalb der Schulformen die umschriebenen Geschlechterunterschiede zugunsten von Mädchen nur in geringfügigerem Maße, zugunsten von Jungen jedoch in verstärktem Ausmaß vorzuliegen scheinen. In PISA 2000 (Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele et al., 2003) noch ausdrücklich auf Schultypenebene bezogene Analysen stehen im Kontrast zu folgenden PISA-Studien, welche nur noch vereinzelt den Schultyp berücksichtigen. So konnten Stanat & Kunter (2003) mit Daten aus PISA 2000 zeigen, dass unter Berücksichtigung des besuchten Schultyps der Vorsprung im Lesen bei Mädchen deutlich geringer ausfiel als auf der Globalebene; teilweise waren keine Kompetenzunterschiede mehr nachzuweisen. Bei Betrachtung nur einer Schulform, in diesem Falle des Gymnasiums, sanken die Effektstärken ebenso deutlich ab – im Bundesdurchschnitt ergibt sich lediglich noch ein kleiner Effekt zugunsten der Lesekompetenz von Mädchen. Zusätzlich sei an dieser Stelle erwähnt, dass Zimmer, Burba & Rost (2004) in allen Kompetenzbereichen *Vorteile* für Jungen unter Berücksichtigung der besuchten Schulform finden. Es ergibt sich bei genauer Betrachtung der erhobenen Kompetenzbereiche der internationalen Schulleistungsstudien ein sehr differenziertes Bild: Wenngleich auf Globalebene ein deutlicher Nachteil für Jungen in der

Lesekompetenz festgestellt werden kann, werden gefundene Effekte hinsichtlich ihrer Stärke deutlich geringer, wenn man die Ebene der Schulform einbezieht und darüber hinaus Kompetenzvorsprünge von Jungen in Mathematik noch deutlicher.

Zusätzlich zu berücksichtigen bleibt als zweiter, zentraler Indikator die Geschlechterverteilung auf bestimmte Schulformen, einhergehend mit den erworbenen Abschlusszertifikaten. Bei Betrachtung des Zertifikatserwerbs im Schuljahr 2013/14 wird deutlich, dass ein vergleichsweise wesentlich größerer Anteil der männlichen Schüler (6.2%) die Schule ohne Abschluss verlässt (Mädchen: 4.2%). Auch beim niedrigsten deutschen Schulabschluss, dem Hauptschulabschluss, stellen Jungen mit einem Anteil von 19.5% aller Jungen eine deutlich größere Gruppe dar als Mädchen mit 14.2%. Während auf Realschulen ein ausgeglichenes Bild vorherrscht (42.2% Jungen; 42.1% Mädchen), zeigt sich auf der höchsten deutschen Schulform, dem Gymnasium, wiederum ein Unterschied zugunsten von Mädchen, welche mit 39.5% aller Schülerinnen einen größeren Anteil der Abiturienten/-innen stellen als Jungen, von denen nur 32.1% das Abitur erlangen (Statistisches Bundesamt, 2014a). Entsprechende Geschlechtsunterschiede lassen sich bei der Betrachtung der Schulbesuchszahlen auf den jeweiligen Schulformen feststellen. Zusätzlich zeigt sich hier, dass im Vergleich zu Mädchen ca. doppelt so viele Jungen eine Förderschule besuchen (220 469 Jungen zu 122 874 Mädchen im Schuljahr 2013/14, ebd.). Weiter sei darauf hingewiesen, dass in Anlehnung an den nationalen Bildungsbericht keine interaktiven Effekte zwischen Geschlecht und Migrationsstatus nachgewiesen werden konnten (Konsortium Bildungsberichterstattung, 2006), weshalb nicht davon auszugehen ist, dass ein geringerer schulischer Erfolg nur auf Jungen dieser sozialen Gruppe zurückzuführen ist.

Die Aktualität dieses Befundmusters wird deutlich, wenn vergleichend frühere Alterskohorten hinzugezogen werden. Im Großen und Ganzen haben sich geschlechtsspezifische Unterschiede in der Bildungsbeteiligung und im Bildungserfolg in Deutschland in den letzten Jahrzehnten umgekehrt. Weitere Daten aus dem Mikrozensus 2013 (Statistisches Bundesamt, 2014b) zeigen, dass der Anteil von Frauen mit dem höchsten Schulabschluss (Abitur) in der Kohorte heutiger 60- bis 65-jähriger Frauen (17.2%) deutlich unter der Abschlussquote der Männer (26.9%) liegt. Auf dem niedrigsten Schulzweig (Hauptschule/Volksschule) erwarben wiederum mehr Frauen ein Abschlusszertifikat (42.0% der Männer, 45.2% der Frauen). In der Alterskohorte der heute 40- bis 45-Jährigen hat sich die Quote an Männern mit Abitur im Zuge der Bildungsexpansion geringfügig gesteigert, Frauen hingegen erreichten deutlich häufiger ein Abitur als in früheren Kohorten. Jedoch ist festzustellen, dass noch immer mehr Männer ein Abitur erwarben, wenngleich der Niveauunterschied zwischen Männern und Frauen deutlich gesunken ist. Wie auch Helbig (2010) feststellt, stieg die Quote an Frauen auf höheren Schulformen kontinuierlich seit den frühen 90er Jahren weiter an und führte in der Alterskohorte der zum Zeitpunkt der Erhebung des Mikrozensus im Jahre 2013 20- bis 25-Jährigen erstmals zu einem ent-

gegengesetzten Muster: 20.7% der männlichen Befragten und nur 13.5% der weiblichen verließen die Schule mit einem Hauptschulabschluss, dem niedrigsten deutschen Schulzertifikat, wohingegen 42.1% der Männer und 50.5% der Frauen dieser Kohorte die Schule mit einem Abiturzeugnis verließen (Statistisches Bundesamt, 2014b). Rau (2013) überprüfte die Effektstärke der zugrundeliegenden Verteilung und kam zu dem Schluss, dass mit $\varphi_{Cramer} = .112$ ein nur als klein einzuschätzender Befund vorliegt. In Anlehnung an die o.g. Befunde aus den PISA-Studien wird nichtsdestotrotz ergänzend deutlich, dass durch die Überrepräsentanz von Jungen auf vergleichsweise niedrigeren Schulformen ein Kompetenzgefälle zwischen Jungen und Mädchen entstehen kann, da erwartungskonform auf niedrigeren Schulformen sowohl die erreichten Kompetenzpunkte als auch die erreichten Kompetenzstufen deutlich von denen in höheren Schulformen abweichen.

Abschließend lassen sich auch hinsichtlich anderer Bildungsindikatoren, wie z.B. bei der Notenvergabe (Helbig, 2010; Neugebauer, 2011) wenn überhaupt nur vergleichsweise sehr geringe Bildungsdisparitäten zwischen den Geschlechtern feststellen. Es kann somit festgehalten werden, dass sich die umfassende Benachteiligung von Jungen, wie sie derzeit medienwirksam propagiert wird (vgl. folgender Abschnitt) als empirisch nicht haltbar erwiesen hat. Es soll nicht bestritten werden, dass Benachteiligungen von Jungen festgestellt werden können, jedoch sind die Effektgrößen der festgestellten Befunde häufig nur klein. Diese Befunde sprechen dafür, dass auch in der Zukunft Anstrengungen notwendig sein werden, Geschlechtergerechtigkeit in der Schule für Jungen *und* Mädchen zu gewährleisten (u.a. Cornelißen, 2004).

Weitere Aspekte von Jungenverhalten mit Relevanz für den schulischen Kontext liefern Befunde zu Unterrichtsstörungen und Aggressivität, welche einen wichtigen Indikator für deviantes Verhalten in der Schule darstellen. Hierbei zeigt sich, dass Jungen mit 7.9% deutlich häufiger als Mädchen mit 1.8% wegen Störungen des Sozialverhaltens und hyperkinetischen Verhaltensstörungen behandelt werden (Schlack, Hölling, Kurth & Huss, 2007) und diese Schülerinnen und Schüler von Lehrkräften als eine erhebliche Belastung erlebt werden (Budnik & Skale, 2007). Per Selbstbericht schätzen sich Jungen im Vergleich zu Mädchen als aggressiver ein und werden auch von den Lehrkräften als physisch aggressiver wahrgenommen (Krahé, Busching & Möller, 2012). Die Aggressivität von Jungen in der Schule wird durch ihren häufig direktverletzenden Charakter eher auffällig als sozial-indirekte Aggressivität von Mädchen (vgl. Eagly & Steffen, 1986; Hyde, 1984). Auch hinsichtlich Stresserleben und psychischer Gesundheit zeichnen sich Geschlechterunterschiede ab. So weist Paulus (2003) auf ein zweifach höheres Risiko für Jungen im Kindesalter hin, an einer psychischen Störung zu erkranken. Stresserleben in der Schule empfinden Jungen und Mädchen zwar gleichermaßen stark, dennoch nehmen Jungen deutlich weniger die Möglichkeit zur Stressverarbeitung, z.B. durch Gespräche über Probleme, wahr (Krebs, 2002). Insgesamt geben diese Befunde weitere Hinweise, dass möglicherweise

Jungen in ihrer Leistungsfähigkeit und ihrem Kompetenzerleben in der Schule beeinträchtigt sind.

Mediale Aufbereitung der „Jungenproblematik“. Die mediale Berichterstattung greift das Thema von Jungen als "Bildungsverlierern" bereits seit längerer Zeit auf und es ist eine Wende von der ehemals dominanten Debatte um benachteiligte Mädchen hin zur Betonung der schulischen Situation von Jungen festzustellen. Erste große, mediale Aufmerksamkeit erreichte die am 17.05.2004 veröffentlichte Ausgabe des Nachrichtenmagazins „Spiegel“, welche mit der Titelgeschichte den Diskurs um „Schlaue Mädchen – dumme Jungen?“ nachhaltig geprägt hat. In vielen anderen Zeitungen und Zeitschriften wurde das Thema in der Vergangenheit ebenso umfassend bearbeitet. Eine Auswahl der letzten Jahre stellen folgende Publikationen dar. Tageszeitungen: *Die Welt*, 24.04.2013: „Mädchen hängen Jungen immer deutlicher ab“; *Frankfurter Allgemeine Zeitung*, 11.04.2013: „Faule, fahrigere Jungen“; *Süddeutsche Zeitung*, 07.01.2013: „Wieso Jungen schlechtere Noten bekommen“; *Die Welt*, 13.07.2012: "Erfolgsrezept Mädchen"; *Die Welt*, 02.07.2012: "Jungs sind die Verlierer"; *Frankfurter Rundschau*, 01.03.2011: "Das Problem der Jungen ist, dass sie Jungen sind"; *Berliner Zeitung*, 16.12.2010: "Das Jungs-Problem"; *Die Zeit*, 05.08.2010: "Jungs sind so!"; Magazine und Zeitschriften: *Süddeutsche Zeitung, Magazin Kinderleben*, 11.11.2008: „Sind Mädchen die neuen Jungs?“; *Psychologie heute*, 14.02.2008: „Männer: neue Erkenntnisse über ein schwaches Geschlecht“; *Spiegel Special*, 26.02.2008: „Das starke Geschlecht: Was Frauen erfolgreich macht“; *Wirtschaftswoche*, 29.10.2007: „Wie die Schulen unsere Jungs verdummen“; *GEO*, 03.03.2003: „Jungs – werden sie die Sorgenkinder unserer Gesellschaft?“; Fernsehen: *3sat*, 25.03.2009, 15.07.2010: „Mehr Wissen über Jungs. Sind sie das neue schwache Geschlecht?“; Radio: *Deutschlandradio*, 02.04.2012: „Debatte: Jungen“.

Auffallend beim Review der Medienbeiträge ist, dass hier von einer generellen und allumfassenden Benachteiligung der Jungen ausgegangen wird, die empirisch anhand der eingangs aufgezeigten Kennzahlen nicht haltbar scheint. Zwar geben die Kompetenzgefälle im Lesen und Unterschiede beim Besuch höherer Schulformen sowie beim Erwerb von Abschlusszertifikaten Anlass zur Frage, ob Jungen im Bildungsbereich mit Schwierigkeiten konfrontiert werden, die ihren Schulerfolg erschweren. Nichtsdestotrotz kann von einer generellen Benachteiligung von Jungen nicht ausgegangen werden, da, wie im vorhergehenden Abschnitt deutlich wurde, auf Ebene des Schulunterrichts in einigen Fächern noch ein Vorteil für Jungen oder zumindest eine gleiche Leistungsfähigkeit von Jungen und Mädchen gefunden werden kann. Dennoch bleiben Medien ein bedeutsamer Bestandteil der alltäglichen Lebenswelt und tragen maßgeblich zur öffentlichen Meinungsbildung bei, indem sie Sachverhalte interpretieren, bewerten und öffentlich darstellen (Eilders, Neidhardt & Pfetsch, 1997). Demzufolge kann die mediale Darstellung trotz ihrer Konzentration auf nur einzelne Aspekte und überzeichneten Darstellung der Problematik einerseits für Jungen zur Bedrohung erwachsen und andererseits ebenso das Problemver-

ständnis in der breiten Öffentlichkeit maßgeblich mit prägen. Es gilt somit, einen potentiellen Einfluss der medialen Berichterstattung zur Thematik „Jungen als Bildungsverlierer“ bei der Bewertung aktueller Befunde zu berücksichtigen.

Aufbauend auf dieser Betrachtung von zentralen Kennziffern zum Bildungserfolg von Jungen und Mädchen im deutschen Schulsystem und einer Betrachtung des medialen Echos auf Geschlechterdisparitäten im Kompetenz- und Zertifikatserwerb soll im Folgenden darauf eingegangen werden, welche theoretischen Erklärungsmöglichkeiten für die Entstehung und Aufrechterhaltung von Geschlechterunterschieden im Allgemeinen sowie geschlechterbezogenen Bildungsungleichheiten im Speziellen existieren.

1.2. Entstehung von Geschlechterunterschieden

Unterschiede im sozialen Verhalten von Frauen und Männern erklären Eagly (1987) sowie Eagly, Wood & Diekmann (2000) mit ihrer sozialen Rollentheorie. Es wird davon ausgegangen, dass die gesellschaftliche Arbeitsteilung sowie die Übernahme verschiedener familiärer Aufgaben zwischen den Geschlechtern sowohl zur Herausbildung von Geschlechterrollenerwartungen als auch zur Entwicklung von typisch geschlechterbehafteten Fähigkeiten und Überzeugungen geführt hat. Dies wiederum wirkt sich schlussendlich in Geschlechterunterschieden im sozialen Verhalten aus. Eine häufigere Beobachtung von Frauen in sozialen Rollen, in welchen das Wohlbefinden von Anderen und fürsorgliches Verhalten im Vordergrund stehen, führt zu der Assoziation, dass Frauen in einem höheren Ausmaß als Männer durch mit diesen Verhaltensweisen verbundene Eigenschaften beschrieben werden können. Typisch männliches Rollenverhalten betont wesentlich häufiger instrumentelle, führende Eigenschaften (Diekmann & Goodfriend, 2006; Wood & Eagly, 2002). Dies führt - vereinfachend - zur Charakterisierung von Männern als „*agentic*“ und von Frauen als auf „*communion*“ hin orientiert. Auch Kinder lernen aufbauend auf diesen Kategorien rasch die soziale Bedeutung von Geschlecht und beginnen frühzeitig, diese Rollenmodelle zu imitieren, sichtbar in geschlechterspezifischen Verhaltensweisen und Eigenschaften (Wood & Eagly, 2002).

Im Falle, dass sich die Arbeitsteilung zwischen den Geschlechtern im Zuge einer gesellschaftlichen Modernisierung angleicht, sollte es zu einer entsprechenden Anpassung der Geschlechterrollen kommen: Es wäre in diesem Falle zu erwarten, dass Männer vermehrt mit derzeit typischerweise feminin konnotierten Eigenschaften und im Gegenzug Frauen des Öfteren mit derzeit als maskulin geltenden Eigenschaften beschrieben werden (Eagly & Wood, 1999). Während der Status und die typischen Charakteristiken von Männern in der heutigen, arbeitsteiligen Gesellschaft kaum Veränderungen erfahren haben, hat sich das gesellschaftliche Bild von Frauen geändert, denn mehr Frauen gehen einer bezahlten Tätigkeit nach. Allerdings sind die

meisten Berufe von Frauen noch immer deutlich geschlechtskonnotiert (Büroarbeit, soziale und pflegende sowie Assistenzberufe). Es bestehen immer noch Unterschiede in der Vergütung von Frauen und Männern trotz gleicher Arbeitsleistung und unbezahlte Haus- und Familienarbeit ist noch immer fest in weiblicher Hand (Heisig, 2011; Leuze & Strauß, 2009). Nichtsdestotrotz kann das vermehrte Teilnehmen von Frauen am Arbeitsleben bereits zu einer Veränderung der sozialen Rollenerwartung geführt haben, wofür metaanalytische Studien von Twenge (1997, 2001) sprechen, nach denen Frauen vermehrt mit mehr maskulineren Eigenschaften beschrieben werden, ohne dass die Zuschreibung femininer Eigenschaften abgenommen hätte. Diese Ergebnisse können als wichtiger Hinweis auf die Validität der sozialen Rollentheorie gewertet werden, da sich in der Tat Änderungen in der Arbeitsteilung zwischen den Geschlechtern auch in geänderten sozialen Rollenerwartungen manifestieren.

Wood & Eagly (2002) postulieren abschließend zwei Wirkrichtungen, über welche das Verhalten und Erleben von Menschen durch Geschlechterrollen beeinflusst werden kann und zwar a) durch die Internalisierung / Übernahme der Geschlechterrollen sowie der damit verbundenen Intention, sich geschlechterrollenkonform zu verhalten sowie b) über einen gewissen normativen Druck, der von den Geschlechterrollenerwartungen der sozialen Umwelt ausgeht. Beide Mechanismen können im Kontext der vorliegenden Arbeit einen wichtigen Beitrag zur Erklärung von Jungenverhalten in der Schule und von Bildungsbeteiligung und -erfolg von Jungen leisten.

Eine weitere, stark kognitiv orientierte Theorie zur Erklärung der Entstehung von Geschlechterunterschieden stellt die Gender Schema Theorie (Bem, 1981, 1983) dar. Die Gender Schema Theorie untersucht, wie Kinder Informationen aus der Umwelt entnehmen und diese Informationen auf sich selbst und soziale Gruppen übertragen. Zentraler Inhalt ist des Weiteren, wie Kinder ihre akquirierten Wissensstrukturen nutzen, um Wahrnehmungen zu organisieren. In einem letzten Schritt betrachtet die Theorie, wie die Wissensstrukturen Prozesse wie Wahrnehmung, Aufmerksamkeit und Verhalten beeinflussen (Martin, Ruble & Szkrybalo, 2002). Der Darstellung von Martin et al. (2002) folgend werden *Gender Schemata* als dynamische Wissensrepräsentationen beschrieben, die sich auf die altersbedingte Entwicklung von Geschlechterbildern beziehen und aus der Interaktion von Individuum und Umwelt resultieren. Die Wahrnehmung von Geschlechtern ist von zentraler Bedeutung für die Entwicklung von Geschlechterbildern: Kinder bilden *Gender Schemata* aus, welche Informationen über sich selbst und über typische Eigenschaften der Geschlechter beinhalten. Diese beeinflussen nachhaltig die Informationsverarbeitung und das gezeigte Verhalten. Weiterhin geht die Theorie davon aus, dass Kinder eine aktive Rolle bei der Konstruktion ihrer eigenen, geschlechtsspezifischen Entwicklung einnehmen. Die aktive Rolle des Kindes wird deutlich in der konstruktiven Informationsverarbeitung sowie in der Motivation des Kindes, bestimmte geschlechterbezogene Verhaltens-

weisen anzunehmen, sobald sie anfangen, sich selbst als Junge oder Mädchen wahrzunehmen und nach Unterschieden und Gemeinsamkeiten zwischen den Geschlechtern suchen. Zusätzlich wird die aktive Steuerung der Schemata durch zwei Prozesse betont, welche die *gender schemas* mit dem Verhalten einer Person verbinden: a) Kinder suchen mehr nach Informationen über ihr eigenes Geschlecht und erfahren daher, wie sie sich geschlechterkonform zu verhalten haben und b) Kinder tendieren dazu, sich konform zu den Geschlechternormen zu verhalten, um sich selbst zu definieren. Zusammen mit weiteren Faktoren wie Entwicklungsstand oder Erwartungen signifikanter Dritter wird bestimmt, ob ein Kind sich geschlechterkonform verhält oder nicht. Die Entwicklung von *gender schemas* wird über verschiedene Grundprinzipien des Erwerbs erklärt. Einerseits nutzen Menschen funktional bedeutende Kategorien – wie eben das Geschlecht –, um Informationen aus ihrer Umwelt zu klassifizieren und organisieren. Kinder nehmen bereits geringfügige Hinweisreize wahr, ob und wann ein Geschlecht bestimmte Verhaltensweisen häufiger aufzeigt als das andere. Dies zeigt sich beispielsweise in der Tatsache, dass bereits Kleinkinder Töne und Phoneme wahrnehmen und bestimmten Geschlechtern zuordnen können. Typischerweise werden die Geschlechterkategorien anhand von physischen Merkmalen definiert. Weiterhin nehmen Individuen an, dass Mitgliedern einer Gruppe bestimmte gemeinsame Merkmale zugrunde liegen. Somit wird eine Distinktheit der Gruppen konstruiert, verschiedene Kategorien hängen auch mit unterschiedlichen Merkmalen zusammen. Darauf aufbauend wird der Theorie folgend durch die Kategorisierungsleistung auch das induktive Denken gefördert, was einerseits den Zuwachs der geschlechterschematischen Wissensstrukturen fördert und es andererseits dem Individuum möglich macht, über die momentan verfügbaren Informationen hinausgehende Rückschlüsse zu ziehen. Abschließend zu berücksichtigen bleiben die kognitiven Konsequenzen der Kategorisierung. Die Kategorisierung bewegt Individuen dazu, Gruppenunterschiede zu übertreiben und Gemeinsamkeiten in einer Gruppe hervorzuheben, also eine verstärkte Abgrenzung nach außen zu betreiben (Martin et al., 2002). Insgesamt kommt also der Wahrnehmung und Verarbeitung von geschlechterbezogenen Informationen aus der Umwelt eine herausragende Bedeutung zu. Angewendet auf die Problemstellung dieser Arbeit bedeutet dies: Wenn davon ausgegangen wird, dass am Lernort Schule spezielle Hinweisreize bestehen, nach denen bestimmte Verhaltensweisen „jungentypisch“ oder „mädchentypisch“ sind, kann dies dazu führen, dass entsprechende Verhaltensweisen in ein Geschlechterschema aufgenommen werden und entsprechend das Verhalten der Schülerinnen und Schüler beeinflussen. Durch die aktive Steuerung der Schemata wäre anzunehmen, dass sie sich entsprechend der wahrgenommenen Geschlechternormen verhalten wird. Im Konkreten könnte der Theorie folgend angenommen werden, dass Jungen eine Assoziation von Schule = weiblich (Heyder & Kessels, 2013) wahrnehmen und entsprechende Schemata ausbilden bzw. diese Informationen in bestehende Schemata integrieren, aktiv nach konformen Informationen suchen und schlussendlich zur Aufrechterhaltung der wahrgenommenen Geschlechternorm schulisches Engagement

vermeiden, um sich der eigenen Wahrnehmung folgend konform zu verhalten. Im Kontext von Jungen in der Schule liefert die Theorie von Bem (1981, 1983) somit weitere, wertvolle Erklärungsansätze für Geschlechterdisparitäten im Schulsystem.

1.3. Soziale Konstruktion von Geschlecht in der Schule

Verschiedene Befunde weisen darauf hin, dass Schule im Lebensalltag von Kindern und Jugendlichen allgemein, somit im Speziellen auch für Jungen, einen wichtigen Erfahrungsraum für soziale Interaktionen darstellt (Fend, 2001; Steins, 2005). An diesem Ort scheinen die problematischen Seiten von Jungen besonders stark in den Vordergrund zu treten, denn Schule ist kein geschlechtsneutraler Raum, sondern ein Ort, an welchem die aktive, soziale Konstruktion von Geschlecht eine wichtige Rolle spielt (Faulstich-Wieland, Weber & Willems, 2004). Diese Auffassung von Geschlecht als Produkt einer Darstellungsleistung und nicht als starre Eigenschaft geht auf West & Zimmerman (1987) zurück und kritisiert das bislang vorherrschende Verständnis von Geschlecht als stabil und determiniert. Insbesondere im Hinblick darauf, dass Schule eine zentrale Lebenswelt von Jugendlichen darstellt und ein Ort ist, an dem ein großer Teil des alltäglichen Lebens verbracht wird, überrascht es wenig, dass im Kontext Schule auch Geschlechterrollen und geschlechtliche Identitätsprozesse von Jungen und Mädchen erworben werden. Connell (2000) schreibt der Schule sogar die Funktion einer institutionalisierten Vermittlung von Männlichkeitsvorstellungen zu, ähnlich beschreibt Budde (2005): „*die Schule als Feld zur Herstellung von Männlichkeit*“ (S. 9).

Der Schule kommt bei der Konstruktion von Geschlecht somit eine besonders wichtige Bedeutung in sozialen Interaktionen der Schüler/-innen zu. Von klein auf erlernt und zumeist kontinuierlich verstärkt haben Schüler/-innen ein breites Spektrum an Vorstellungen und Erwartungen entwickelt, wie das eigene und das andere Geschlecht zu sein oder sich zu verhalten haben (Prentice & Carrenza, 2002). Der eingeführte Begriff des „*doing gender*“ zielt dabei darauf ab, die Geschlechtszugehörigkeit nicht als unveränderliche Eigenschaft oder als Merkmal von Individuen zu begreifen, sondern vielmehr die sozialen Prozesse in den Fokus zu rücken, in welchen Geschlecht als Unterscheidungskriterium herangezogen wird (Gildemeister, 2004). Wie im Alltag, so auch in der Schule konstruiert sich Geschlecht aus sozialer Interaktion, in der einerseits Geschlecht inszeniert und andererseits diese Inszenierung anerkannt wird. Über verschiedene, schulbezogene Verhaltensweisen lässt sich im Kontext Schule ebenfalls Geschlecht inszenieren und konstruieren. Dabei fällt auf, dass „typisches Jungenverhalten“ des Öfteren eher mit lern- und bildungshinderlichem Verhalten (für einen Überblick aus Lehrerperspektive: Budde, Scholand & Faulstich-Wieland, 2008) verbunden zu sein scheint, wohingegen ein „typisches Mädchen“ durch bessere affektive (Segeritz, Stanat & Walter, 2010) und kognitive Einstellungen

zur Schule, mehr Selbstdisziplin und stärker selbstgesteuerte Arbeitsweisen (Duckworth & Seligman, 2006) und ein generell gründlicheres Lern- und Arbeitsverhalten (z.B. Hausaufgaben; Spiel, Wagner & Fellner, 2002) auffällt und damit diese Verhaltenseigenschaften eher feminin konnotiert werden, oder – in der Terminologie des „*doing gender*“ ausgedrückt – für die Inszenierung von Geschlecht in der Schule relevant werden und zur Inszenierung herangezogen werden können. Kessels & Hannover (2004, 2006) nehmen an, dass die Entwicklung schulischer Interessen (welche Fächer werden bevorzugt, welche abgelehnt – bis hin zum generellen Interesse für alle schulischen Inhalte) mit der Identitätsentwicklung im Jugendalter zusammenhängt. Jugendliche nutzen das Angebot der Schule, um bestimmte Aspekte der eigenen Identität zu entwickeln und zu inszenieren. In dem Maße, wie sich ein Schüler / eine Schülerin in bestimmten Schulfächern – oder in der Schule überhaupt – engagiert, handelt es sich keinesfalls um reine Akquise von Fachkenntnissen und Kompetenzen, sondern gleichfalls auch um die der Inszenierung der den entsprechenden Inhaltsdomänen oder Verhaltensweisen anhaftenden, sozialen Bedeutungen. Es existieren sozial geteilte Annahmen über die Bedeutung des Engagements in einem bestimmten schulischen Bereich oder aber in der Schule generell. Diese werden Teil des Selbst des Schülers bzw. der Schülerin integriert, wenn ein entsprechendes Engagement gezeigt wird (Kessels & Hannover, 2004, 2006). Mit diesem Erklärungsansatz kann nicht nur verdeutlicht werden, warum sich Mädchen selten den mathematisch-naturwissenschaftlichen Bereichen zuwenden, sondern auch, warum es für Jungen aus dem Blickwinkel der Entwicklung der Geschlechtsidentität funktional sein könnte, schulisches Engagement zu vermeiden: Wird der Bereich der Schule (oder aber nur ein Teilbereich, Fach etc.) von Jungen wie oben beschrieben als weiblich konnotiert wahrgenommen, würde ein entsprechendes Engagement von Jungen deren Selbstdefinition gefährden. Nur durch das Vermeiden von (fachspezifischem oder generellem) schulischem Engagement kann die Selbstdefinition stabilisiert und kann eine gesicherte Maskulinität zur Schau gestellt werden. Schulisches Engagement birgt so betrachtet für Jungen psychologische Kosten und gefährdet potentiell ihre Identität als „richtiger Junge“ (Hannover & Kessels, 2011).

Zahlreiche, zum bisherigen Zeitpunkt allerdings eher qualitativ ausgerichtete Studien, haben untersucht, wie sich „*doing gender*“ im aktuellen Schulalltag darstellt und beziehen vermehrt auch Lehrkräfte als aktive Mitkonstrukteure von Geschlecht als relevanter Kategorie im Schulalltag mit ein. Michalek (2009) präsentierte Ergebnisse aus Gruppendiskussionen zweier Grundschulgruppen und stellte fest, dass trotz der Herstellung einer gemeinsamen Basis des „Junge seins“ zwischen den Gruppen deutlich divergierende Geschlechtervorstellungen berichtet wurden. In einer der untersuchten Gruppen wurde mit der Geschlechterkategorie weiblich ein vollständiger Gegenpol zum Junge sein konstituiert, stereotype Ansichten über eine generelle (auch schulische) Überlegenheit der Jungengruppe gegenüber der Gruppe der Mädchen über-

wogen in den Äußerungen sehr stark (*ingroup bias*), ebenso das Bild des „Helden“, der viel ertragen kann und sich aufopfert. Für die Darstellung der Überlegenheit der Eigengruppe wurden teilweise sogar eher weiblich konnotierte Eigenschaften herangezogen. Diese Jungengruppe war in ihren Äußerungen äußerst homogen und dichotomisierte enorm. Der körperlichen Überlegenheit und Sportlichkeit kommen *sensu* Frosh, Phoenix & Pattman (2003, S. 100) auch in dieser befragten Gruppe eine immense Bedeutung für die soziale Konstruktion von Männlichkeit zu. Die ergänzend untersuchte Gruppe zeigte diese Zuschreibungen und Typisierungen hingegen kaum und überraschte durch sehr heterogene Sichtweisen und stand in starkem Kontrast zur eingangs geschilderten Gruppe. Auch hinsichtlich der Thematik Schule und Schulalltag unterschieden sich beide Gruppen deutlich voneinander. Während in einer Jungengruppe in den Gesprächspassagen zum Thema Schule eine deutliche Distanz zu diesem Kontext zum Tragen kam, wies die andere Gruppe hier wieder deutlich differenziertere Aussagen auf, welche keinen eindeutig negativen oder positiven Bezug zur Schule erkennen ließen, sondern vielmehr die individuelle Positionierung zu bestimmten Inhalten, Schulfächern und deren Bedeutung zum Tragen kam. Inwieweit durch die Positionierung zu gewissen Schulfächern „*doing gender*“ wiederum durch geschlechterrollenkonforme Interessen und Fachpräferenzen hervortritt, ist leider kein expliziter Bestandteil dieser Untersuchung. Inwieweit die gefundenen Differenzen auf Merkmale der Gruppe oder ihrer Mitglieder zurückzuführen sind, kann anhand dieser Studie ebenfalls zunächst nicht beantwortet werden. Zusammenfassend kann als zentrale Aussage abgeleitet werden, dass es auf Grundlage der vorliegenden Ergebnisse schwierig ist, im Allgemeinen von „den“ Jungen zu sprechen, da die Befunde deutliche Divergenzen zwischen zwei exemplarischen Jungengruppen indizieren. Es kann abschließend vermutet werden, dass verschiedenartige Orientierungen des „*doing gender*“ allerdings mit ebenso verschiedenartigen Einstellungen und Orientierungen in Bezug auf die Schule zusammenlaufen, eine Orientierung auf Heterogenität scheint anhand der vorliegenden Befunde mit förderlicheren Einstellungen gegenüber Leistung und Schule einherzugehen.

Faulstich-Wieland et al. (2004) gingen in einer vierjährigen Längsschnittstudie mittels ethnografischer Beobachtung und einiger standardisierter Fragebögen der Fragestellung nach, wie Lehrkräfte und Schüler/-innen in verschiedenartig konfigurierten Schulklassen durch Interaktionen in einer Vielzahl an Schulfächern Geschlecht als soziale Kategorie herstellten und in welchen Situationen geschlechterstereotype Inszenierungen zu einer sog. „*ruhenden Ressource*“ gerieten. Es konnte festgestellt werden, dass Lehrer/-innen das Geschlecht wesentlich stärker dramatisierten, als es nach reiner Betrachtung der Schüler/-inneninteraktion zu erwarten gewesen wäre. Neben zahlreichen Interaktionen und „*Spielräumen*“, in welchen Geschlecht als soziale Kategorie bewusst ausgespielt wurde (z.B. zur Aushandlung eines Status), konnten interessanterweise auch zahlreiche Interaktionen mit „*undoing gender*“, d.h. ohne die Aktivierung von Ge-

schlecht als soziale Kategorie gefunden werden. Die These, dass Inszenierungen von Geschlecht insbesondere in Klassengruppen mit Minderheitenstatus (z.B. wenn wesentlich mehr Mädchen die Klasse besuchen als Jungen) verstärkt beobachtet werden können, konnte nicht bestätigt werden. Hieraus wird deutlich, dass Geschlecht als soziale Kategorie und „*doing gender*“ sehr wohl immer verfügbar ist, aber keineswegs in jeder Interaktion herangezogen wurde. Andere Kategorien, wie etwa das Alter („*doing adult*“), waren in einigen Interaktionen eine deutlich bedeutsamere Kategorie. Als pädagogisch notwendige Konsequenz wurde abgeleitet, dass Mädchen nicht im Allgemeinen auf Grundlage der Zuschreibung des Geschlechts und damit verbundener Rolleneigenschaften als benachteiligt wahrgenommen und Jungen nicht auf ihr Störungspotenzial reduziert werden sollten.

Budde et al. (2008) stellten in einer Interviewstudie hingegen fest, dass eine von Michalek (2009) geforderte Akzeptanz von Differenzen bei der Definition von „Junge sein“ im Klassenzimmer noch nicht angekommen zu sein scheint: Innerhalb der Lehrer-Schüler-Interaktion erfahren Jungen häufiger Ablehnung und Mädchen wiederum des Öfteren Wertschätzung, was sich weiterführend auf die Positionierung im Schulkontext auswirkt. Während Jungen von Lehrer/-innen eher als negativ-auffällig eingeschätzt wurden, wurden Mädchen entgegengesetzt als positiv-unauffällig charakterisiert. In der Benotung zeigte sich weiterhin, dass insbesondere die auffällig-negativen Jungen (also ein Großteil der männlichen Stichprobe) deutlich schlechtere Noten erzielten, als es nach einer standardisiert gemessenen Testleistung (HST 4/5) zu erwarten gewesen wäre. Dies kann als Hinweis darauf interpretiert werden, dass vor dem Hintergrund einer differenziell schlechteren Benotung aufgrund des Geschlechts bei gleichen Kompetenzen (vgl. Kap. 1.4.) Jungen ihr kognitives Potential weniger ausnutzen, weil sie schulische Leistung als schlechter vereinbar mit ihrer Geschlechterrollenerwartung wahrnehmen. Es wurde resümiert, dass insbesondere soziale Aspekte im Kontext Schule für Jungen zum Hindernis werden. Während positive soziale Verhaltensweisen von Mädchen gezeigt und ihnen auch zugetraut werden, zeigt sich bei Jungen ein entgegengesetztes Bild. Von ihnen wird störendes Verhalten geradezu erwartet, dementsprechend werden ihre Handlungen und Beiträge auch interpretiert. Dies gipfelt in dem Dilemma, dass von Jungen zwar adäquates Verhalten in der Schule erwartet wird, ihnen aber andererseits per Geschlecht und den damit verbundenen Rollenerwartungen genau die Kompetenz zu sozial angemessenem Verhalten abgesprochen wird.

Parker und Rennie (2002) untersuchten in einer interviewbasierten Studie im Rahmen des australischen „Single-Sex Education Pilot Project“ (SSEPP) Schülerinnen und Schülern der Jahrgangsstufen 8-10 mit sowohl mono- als auch koedukativen Schul- und Lernerfahrungen. Ebenso wurde die Perspektive der Lehrer/-innen als zusätzliches Kriterium mit herangezogen. Auf Ebene der Lehrkräfte konnten die Autoren zeigen, dass die Lehrpersonen für Mädchen monoedukative Klassen bevorzugten, wohingegen für Jungen das Meinungsbild deutlich zur Ko-

edukation tendierte. Ausschlaggebend für diesen Befund war u.a. die Tatsache, dass Mädchenklassen als einfacher im Klassenmanagement im Vergleich zu koedukativen oder reinen Jungenklassen wahrgenommen wurden. Es wurde somit *qua Geschlecht* eine soziale Wirklichkeit konstruiert, in welcher Mädchen sein auch von Seiten der Lehrpersonen mit unproblematischem Lernverhalten und leichterem Klassenmanagement assoziiert wurde. Auf Ebene der Schüler/-innen präferierten Mädchen häufiger monoedukativen Unterricht, bei Jungen waren die Präferenzen eher ausgeglichen. Während Mädchen von gleichgeschlechtlichem Unterricht zu profitieren scheinen, indem sie z.B. aktiver und extravertierter am Unterricht teilnahmen und mehr Interaktionen mit der Lehrkraft zeigten, fanden sich bei Jungen derartige vorteilhafte Effekte in geschlechtshomogenen Klassen nicht. Es wurden hierbei durch die befragten Jungen nur außerschulische „Vorteile“ betont, wie die Möglichkeit, sexistischer sein und sich einfacher über „Jungenthemen“ austauschen zu können. Holz-Ebeling, Grätz-Tümmers & Schwarz (2000) fanden an einer deutschen Stichprobe ergänzend Hinweise darauf, dass Jungen nicht notwendigerweise von Koedukation profitieren und wiesen auf die Potentiale monoedukativer Beschulung von Jungen hin, wenngleich die Gemeinsamkeiten zwischen beiden Jungengruppen (Mono- und Koedukation) deutlich gegenüber den wenigen gefundenen Unterschieden überwogen. Beachtenswert bleibt, dass bei „harten“ Merkmalen wie Leistung (gemessen wurden allerdings nur Physik- und Englischleistung) tatsächlich die koedukative Schulform mit Vorteilen für Jungen verbunden war, wohingegen „weichere“ Kriterien der sozio-emotionalen Befindlichkeit überwiegend günstiger für Jungen an monoedukativen Schulen ausfielen. Auch zeigten Jungen der monoedukativen Schulform tendenziell stereotypere Fächerwahlen (bspw. Leistungskurse: 55% Mathematik, 28% Englisch; koedukative Schulform: 29% Mathematik, 41% Englisch). Zusätzlich stellten Parker & Rennie (2002) fest, dass die beteiligten Lehrkräfte ihr eigenes Verhalten in koedukativen Klassen wesentlich tiefgreifender reflektieren konnten: Die eingesetzten Strategien im Schulalltag traten hier deutlicher im Bewusstsein der Lehrenden hervor und führten zu verbesserten Erkenntnissen darüber, welche unterschiedlichen Verhaltensweisen gegenüber Mädchen und Jungen eingesetzt und gezeigt wurden und wie diese sich wiederum sowohl auf die Konstruktion der geschlechtlichen Identität der Schüler/-innen als auch auf die Geschlechtergerechtigkeit in der Schule auswirken.

Heyder & Kessels (2012) untersuchten in einer Befragung von Gymnasiallehrkräften den Einfluss von männlich und weiblich inszenierten Verhaltensweisen auf die Zuschreibung von lernförderlichen und lernhinderlichen Eigenschaften. Hierzu wurden in einem ersten Schritt 82 Verhaltensweisen von $N = 104$ Lehrkräften hinsichtlich des Inszenierungspotentials und von $N = 82$ Lehrer/-innen hinsichtlich der sozialen Erwünschtheit bewertet. Bei der Auswahl von geeigneten Verhaltensweisen zur Erstellung des Studienmaterials war es neben der Eignung zur Inszenierung von Männlichkeit oder Weiblichkeit von Bedeutung, dass die Verhaltensweisen sozi-

al gleich (un)erwünscht waren und dass sich in den einzelnen Bewertungen keine Geschlechterunterschiede beim Inszenierungspotential und der sozialer Erwünschtheit nachweisen ließen. Zur Anfertigung der abhängigen Variablen wurden mit $N = 61$ Lehrkräften insgesamt 61 lernrelevante Verhaltensweisen als lernförderlich oder lernhinderlich klassifiziert. Aufgrund der Ergebnisse aus diesen Vorstudien wurden vier Vignetten erarbeitet, in welchen ein Mädchen bzw. ein Junge geschlechterneutrales vs. stark geschlechtsinszenierendes Verhalten zeigte. Es wurden $N = 104$ Gymnasiallehrkräften zufällig eine der vier Vignetten vorgelegt, die im Anschluss beurteilen sollten, wie wahrscheinlich die Person die vorgetesteten lernhinderlichen oder lernförderlichen Verhaltensweisen zeigt. Im Ergebnis zeigte sich bei der Zuschreibung von lernförderlichen Verhaltensweisen ein Haupteffekt des Geschlechts der Stimulusperson in Richtung einer geringeren Anzahl lernförderlicher Verhaltensweisen, die Jungen zugeschrieben wurden, unabhängig davon, wie stark sie tatsächlich das eigene Geschlecht inszenierten. Ergänzend fand sich ein Haupteffekt der Geschlechtsinszenierung: Unabhängig vom Geschlecht der Stimulusperson wurden hier Personen, die ihr Geschlecht stark inszenierten, mit vergleichsweise weniger lernförderlichen Verhaltensweisen charakterisiert. Fokussiert man auf die lernhinderlichen Verhaltensweisen, zeigte sich ein spiegelbildliches Ergebnismuster: Männlichen Stimuluspersonen wurden generell mehr lernhinderliche Eigenschaften zugeschrieben als weiblichen. Zusätzlich wurden Stimuluspersonen mit starker Geschlechtsinszenierung signifikant mehr lernhinderliche Verhaltensweisen zugeschrieben. Zusammenfassend kann hieraus abgeleitet werden, dass Lehrkräfte aufgrund von „*doing gender*“ im schulischen Kontext Jungen *qua* Geschlecht weniger lernförderliche und mehr lernhinderliche Verhaltensweisen unterstellen, wobei die deutliche Inszenierung des eigenen Geschlechts bei Lehrkräften die Erwartung verstärkt, dass der betreffende Lernende lernhinderliche Verhaltensweisen zeigt.

Die in diesem Abschnitt vorgestellten Ergebnisse empirischer Jungenforschung zeigen, dass im schulischen Bereich „*doing gender*“ im Sinne der begrifflichen Definition sowohl aus sozialer Interaktion von Schüler/-innen untereinander entsteht – aber auch die am Bildungsprozess beteiligten Lehrkräfte eine herausragende Rolle bei der Entstehung und Aufrechterhaltung geschlechtsstereotyper Verhaltensweisen einnehmen. Zusammenfassend verdeutlichen die zitierten Studien exemplarisch, dass Lehrpersonen selbst eine aktive Rolle bei der sozialen Konstruktion von Geschlecht im schulischen Kontext zukommt. Auch sie ermöglichen es (zusätzlich), die Bedeutung von Geschlecht salient werden zu lassen und die Ausgestaltung typischer Geschlechtervorstellungen zu produzieren und zu reproduzieren. Damit sind sie weiterhin am Erfolg oder Misserfolg koedukativer Lernsettings, an der Entwicklung geschlechterrollenkonformer Interessen und darüber hinaus an der Entstehung geschlechterdifferenzierter Leistungen mit beteiligt.

1.4. Erklärungsansätze für Geschlechterdisparitäten im Bildungswesen

Im Folgenden soll die teilweise widersprüchliche wissenschaftliche Literatur zu der Frage dargestellt werden, welche Bedingungen in der Schule verantwortlich für den schlechteren Kompetenzerwerb und die mit ansteigender Schulform geringere Bildungsteilhabe von Jungen verantwortlich sind könnten. Das gesamte Kapitel 1.4. folgt in seiner Struktur und den meisten der berichteten Studien der Darstellung von Hannover & Kessels (2011).

Feminisierung von Bildung. Im Primarbereich wird die Erziehung und Ausbildung der Schüler/-innen überwiegend von Frauen geleistet. In Grundschulen lehrten im Schuljahr 2011/12 in Deutschland 87.4% Lehrerinnen. Aber auch an Haupt- und Realschulen waren Frauen mit 62.6% bzw. 65.8% überrepräsentiert. Auch auf der höchsten deutschen Schulform, dem Gymnasium, war der Anteil von Lehrerinnen mit 56.4% leicht höher als der von Lehrern (eigene Berechnungen auf Grundlage eine Datenbankanfrage beim Statistischen Bundesamt). Eine derzeit populäre und bisher am meisten untersuchte Annahme in Bezug auf die schlechtere Bildungsbeteiligung von Jungen besteht darin, dass Jungen aufgrund dieser Mehrbeschäftigung von Frauen im Bildungssektor benachteiligt werden (Helbig, 2010; Quenzel & Hurrelmann, 2010, für einen Überblick). Die weibliche Ausrichtung des Lehrerberufs wird hierbei insbesondere als problematisch angesehen, weil männliche Rollenvorbilder fehlen würden, Lehrerinnen Jungenverhalten eher als abweichend erleben und dementsprechend sanktionieren würden, was wiederum in der empirisch vorzufindenden, schwächeren Bildungsbeteiligung von Jungen resultieren würde. Als weitere Mechanismen werden bessere Bindungen zu gleichgeschlechtlichen Lehrpersonen und dadurch verbesserte Lern- und Arbeitsprozesse in der Schule sowie eine Entfremdung von der Schule für Jungen (durch die Assoziation Schule = weiblich, vgl. auch Heyder & Kessels, 2013) diskutiert. Verschiedene Untersuchungen haben überprüft, ob sich das Geschlecht von Lehrpersonen tatsächlich auf die Interessen- und Leistungsentwicklung der Lernenden auswirkt.

Carrington, Tymms und Merrell (2008) untersuchten anhand von Daten des *Performance Indicators in Primary Schools* (PIPS) Projekts ca. 9 000 Schüler/-innen in 413 Klassen der Jahrgangsstufe 6. Hiervon wurden 113 Klassen von Männern und 300 Klassen von Frauen unterrichtet. Den Schüler/-innen wurde ein Fragebogen vorgelegt, in welchem die Einstellung zur Schule im Allgemeinen, zu Lesen, Mathematik und Naturwissenschaften abgefragt wurden. Zusätzlich wurden Leistungsmaße erhoben. Die Auswertung fand zur Kontrolle der Klassenstrukturen mit einem *multilevel*-Ansatz statt. Die Autoren/-innen zeigten, dass es keine besseren Einstellungen zur Schule und keine besseren Leistungen bei denjenigen Sechstklässler/-innen gab, die von einer Lehrkraft ihres eigenen Geschlechts unterrichtet wurden. Vielmehr profitierten Kinder

beiderlei Geschlechts, die von weiblichen Lehrkräften unterrichtet wurden, diese wiesen insgesamt bessere Einstellungen zur Schule auf.

Schöps, Walter, Zimmer & Prenzel (2006) analysierten anhand der längsschnittlichen PISA-I-Plus-Studie von 2003 Kompetenzunterschiede und Kompetenzentwicklungen von Schülerinnen und Schülern. Auch hier konnte kein Zusammenhang zwischen dem Geschlecht der Lehrperson in der Mathematik und der Kompetenzentwicklung der Schüler/-innen gefunden werden. Daten einer weiteren, internationalen Vergleichsstudie (PIRLS 2006) wiesen ebenso keine konsistenten Zusammenhänge zwischen Lesekompetenzen und Geschlecht der Lehrkraft nach (Mullis et al., 2007).

Helbig (2010) zog zur Beantwortung der Frage, ob Lehrerinnen für den geringeren Bildungserfolg von Jungen verantwortlich sind, Daten aus der ELEMENT-Studie (2003-2005) in den Klassenstufen 4 bis 6 heran. Da hierin keine Angaben über das Geschlecht der Lehrkräfte enthalten waren, wurden diese Daten über die Berliner Senatsverwaltung für Bildung und Forschung für die untersuchten Schulen ergänzt. Als abhängige Variablen dienten die Lese- und Mathematikkompetenz, die entsprechenden Schulnoten sowie die Gymnasialempfehlung. Es konnte gezeigt werden, dass Mädchen zwar geringfügig mit $\beta = .08$ eine bessere Lesekompetenz in Klasse 6 bei einem größeren Anteil weiblicher Lehrkräfte aufwiesen, dass Jungen aber in ihrer Kompetenzentwicklung im Lesen vom Anteil weiblicher Lehrkräfte nicht beeinflusst wurden ($\beta = .02$, n.s.). Ferner erfuhren Jungen in Mathematik eine vergleichsweise geringfügige Benachteiligung hinsichtlich ihrer Benotung. Je höher der Frauenanteil an der Schule war, desto höher fiel die Mathematiknote für Jungen mit $\beta = .01$ aus. Weitere Effekte auf die Benotung wurden nicht statistisch bedeutsam. Besonders zu berücksichtigen bleibt zur Beantwortung der Forschungsfrage, dass sich keine Unterschiede in den erreichten Mathematik- und Lesekompetenzen für Jungen in Abhängigkeit vom Geschlecht der Lehrkraft aufzeigen ließen. Auch für die Empfehlung zur Gymnasialaufbahn war das Geschlecht der Lehrkraft nicht prädiktiv, unter Kontrolle der erreichten Schulnoten in Deutsch und Mathematik. Zwar scheinen Mädchen wie oben beschrieben leicht von gleichgeschlechtlichen Lehrkräften zu profitieren, eine Benachteiligung von Jungen durch Lehrerinnen kann den Ergebnissen von Helbig (2010) folgend aber nicht konstatiert werden.

Auch Neugebauer (2011) unterstützt die Befunde von Helbig, indem auch hier nachgewiesen werden konnte, dass Jungen bezüglich der Gymnasialempfehlung nicht von männlichen Lehrkräften profitierten. Im Gegenteil konnte gezeigt werden, dass für beide Geschlechter ein hoher Anteil von Lehrern eher mit einer zurückhaltenden Praxis bei der Vergabe von Gymnasialempfehlungen verbunden war und beide Geschlechter von einem hohen Lehrerinnenanteil zu profitieren schienen.

Anhand von 964 untersuchten Schüler/-innen prüften Martin & Marsh (2005) den Einfluss des Geschlechts der Lehrkraft auf die akademische Motivation und schulisches Engagement. Hierzu wurde ein geschlechterstereotypes Modell sowie ein geschlechtsinvariantes Modell postuliert und auf empirische Gültigkeit geprüft. Dem geschlechterstereotypen Modell folgend sollten beide Geschlechter von gleichgeschlechtlich unterrichteten Klassen profitieren, wohingegen das geschlechterinvariante Modell von keinem Effekt des Lehrkraftgeschlechts ausging. Die Ergebnisse unterstützten das geschlechterinvariante Modell und zeigten, dass Motivation und Engagement nicht vom Geschlecht der Lehrkraft abzuhängen schien. In einer ergänzenden Untersuchung wurde der Fragestellung nachgegangen, ob Jungen von gleichgeschlechtlichen Lehrkräften profitieren und ob Einflüsse des Klassenklimas vorliegen. Mehrebenenanalysen konnten nur sehr geringe oder gar keine Effekte für diese These nachweisen. Im Allgemeinen wiesen Mädchen zwar generell eine höhere Motivation auf, jedoch unabhängig vom Geschlecht der jeweiligen Lehrkraft. Das Klassenklima als ergänzende Variable konnte hier wesentlich mehr zur Varianzaufklärung beitragen als das Geschlecht der Lehrkraft (vgl. auch Marsh, Martin & Cheng, 2008).

Auch Hadjar & Lupatsch (2011) fanden keinerlei Unterstützung für die These, dass Frauen in Lehrpositionen den Bildungserfolg von Jungen negativ beeinflussen. Neben der wahrgenommenen Fairness und Notengebung wurden weitere Bestimmungsfaktoren für Schulerfolg in diese Studie einbezogen. Die Resultate der Analysen machen deutlich, dass sich Jungen zwar in der Schule etwas mehr diskriminiert bzw. unfair behandelt fühlten. Dies trat allerdings vollkommen unabhängig vom Geschlecht der Lehrperson auf und zeigt sich in gleicher Richtung und Stärke sowohl bei Lehrerinnen als auch bei Lehrern. Auch die Schulnoten in Deutsch und Mathematik unterschieden sich nicht nach dem Geschlecht der Lehrkraft. Ebenso zeigte sich in auf die ersten Analysen aufbauenden Strukturgleichungsmodellen keine andere Befundlage. Auch hier konnte keine direkter Einfluss des Geschlechts der Lehrkraft auf den Schulerfolg in Deutsch und Mathematik nachgewiesen werden. Vielmehr spielten Faktoren wie Fachinteresse, Schulentfremdung und traditionelle Geschlechterrollenkonzepte eine wichtige Rolle für die Erklärung der Befunde, auf welche das Geschlecht der Lehrkraft zwar ebenfalls einen Einfluss hatte, das als solches jedoch nur zu einem vergleichsweise kleinen Teil zur Erklärung des geringeren Schulerfolgs von Jungen beitrug.

Lahelma (2000) stellte in einer qualitativen Interviewstudie mit 90 Befragten fest, dass Geschlecht keine besonders relevante Kategorie darstellte, wenn Schüler/-innen über Lehrkräfte sprachen. Es war – unabhängig vom Geschlecht – für die Schülerinnen und Schüler von zentraler Bedeutung, dass eine Lehrkraft didaktisch kompetent, freundlich und entspannt auftritt. Aber auch Merkmale der Klassen- und Unterrichtsstrukturierung wie „Ordnung halten“ und „Mitarbeit aller Schüler/-innen sicherstellen“ waren von hoher Wichtigkeit für die Befragten. Ge-

schlecht, trotz aktuell brisanter Debatte um fehlende Männer im Bildungswesen, spielte in den Aussagen der Befragten keine bedeutsame Rolle. Die Interviewten gaben ferner an, dass sie keine dringende Notwendigkeit für mehr Lehrer im Bildungssystem im Sinne einer Erhöhung von männlichen Rollenmodellen in der Schule sehen. Explizit auf einen Mangel männlicher Lehrpersonen angesprochen, wurde trotzdem zumeist verneint, dass auf dieser Grundlage Probleme im schulischen Alltag bestehen.

Abschließend sei darauf hingewiesen, dass ergänzend zu den o.g. Befunden auch Untersuchungen existieren, die einen Einfluss des Geschlechts der Lehrkraft nachweisen. So konnten Lörcher & Maier (1999) anhand von Daten der Realschulabschlussprüfung in Baden-Württemberg nachweisen, dass von Lehrerinnen unterrichtete Schülerinnen bessere Mathematiktestergebnisse erreichten. Ebenso wiesen Diefenbach & Klein (2002) nach, dass Schulabschlüsse von Jungen in Bundesländern mit höherem Anteil männlicher Lehrkräfte geringerwertiger ausfielen. Dadurch, dass die auf der Aggregatebene nachgewiesenen Zusammenhänge nicht ebenso auf der Individualebene untersucht und gezeigt wurden, entspricht die abgeleitete Schlussfolgerung allerdings einem ökologischen Fehlschluss (Hannover & Kessels, 2011). Winters, Haight, Swaim & Pickering (2013) konnten mit US-amerikanischen Daten bestätigen, dass Schülerinnen und Schüler von weiblichen Lehrkräften hinsichtlich ihrer Leistungen auf Sekundarniveau profitierten.

In einer weiteren amerikanischen Untersuchung wurde von Dee (2005) mit allerdings nicht mehr als aktuell zu bewertenden Daten aus der NELS:88-Befragung (National Education Longitudinal Study des Jahres 1988) deutlich, dass u.a. geschlechtsbezogene Dynamiken zwischen Schüler/-innen und Lehrkräften deutliche Effekte hervorbrachten, die sich insbesondere auf die lehrkraftbezogene Wahrnehmung der Schüler/-innen-Leistung bezog. So konnte z.B. gezeigt werden, dass Jungen von weiblichen Lehrkräften deutlich häufiger störendes Unterrichtsverhalten vorgeworfen wurde als Schülerinnen. Angemerkt sei allerdings auch, dass äquivalente Effekte auch bei männlichen Lehrkräften in Bezug auf Mädchen aufgezeigt werden konnten, d.h. dass umgekehrt wurden von männlichen Lehrkräften Mädchen häufiger störendes Unterrichtsverhalten vorgeworfen als Jungen. Von einer Benachteiligung von Jungen kann hier also nicht gesprochen werden, vielmehr zeigen die Befunde, dass Geschlecht für die Lehrkräfte eine zentrale Kategorie bei der Wahrnehmung der Lernenden darstellte. In weiteren Analysen (Dee, 2007) wurde darauf aufbauend erneut subsummiert, dass Interaktionen zwischen Schüler/-innen und Lehrkräften (für Mädchen, wie auch für Jungen) einen hohen Geschlechtsbezug aufwiesen und weitere Bereiche jenseits der Leistungserwartungen bedeutsam beeinflussten. Es zeigten sich ferner signifikant bessere Testleistungen und wohlwollendere Bewertungen der Schüler/-innen, wenn sie von einer Lehrkraft des gleichen Geschlechts unterrichtet wurden. Ammermüller & Dolton (2006) weisen diesbezüglich nur auf einen sehr geringen Effekt aus-

schließlich für Jungen in der Mathematikleistung bei Unterrichtung durch einen männlichen Lehrer hin.

Insgesamt sind somit nur wenige empirische Anhaltspunkte für die These ableitbar, dass Frauen im Bildungswesen den schulischen Erfolg von Jungen negativ beeinflussen. Es existieren vergleichsweise zahlreiche Belege gegen die Annahme einer Benachteiligung von Jungen durch Lehrerinnen, wenn derartige Befunde nachgewiesen werden können, dann nur mit vergleichsweise geringer Effektgröße (u.a. Dee, 2007; Helbig, 2010). Vielmehr existieren zusätzliche Befunde, die sogar darauf hindeuten, dass Jungen wie auch Mädchen vom Unterricht bei Lehrerinnen profitieren können (u.a. Carrington et al., 2008; Winters et al., 2013; vgl. Hannover & Kessels, 2011, für einen Überblick)

Geschlechterunterschiede in der Notengebung. Ein weiterer Erklärungsansatz für den geringeren Bildungserfolg von Jungen stellt – der Darstellung von Hannover und Kessels (2011) folgend – die Annahme dar, dass Jungen trotz objektiv gleicher Kompetenzen im Vergleich zu Mädchen generell schlechter benotet werden. Daten der internationalen Schulvergleichsstudien wie der Internationalen Grundschul-Lese-Untersuchung (IGLU 2006; Bos et al., 2007) sowie aus PISA 2006 erlauben eine genauere Betrachtung dieser möglichen Benachteiligung für Jungen. Aus der IGLU-Studie 2006 (Bos et al., 2007) liegen Daten vor, wie viele Kompetenzpunkte ein Kind erreichen musste, das von seiner Lehrkraft eine Empfehlung für eine gymnasiale Schullaufbahn erhalten hatte. Im Bundesdurchschnitt benötigten die in die IGLU-Studie 2006 einbezogenen Jungen für eine gymnasiale Laufbahnempfehlung 585 Kompetenzpunkte, Mädchen hingegen nur 557 Punkte, was ca. dem Kompetenzerwerb eines halben Schuljahrs entspricht. Regressionsanalytisch zeigte sich, dass Mädchen relativ zu Jungen eine 1.25-fache Chance auf eine Gymnasialempfehlung hatten (*cave*: $R^2 = .00$), nach Kontrolle von kognitiven Fähigkeiten und Lesekompetenz konnte noch immer eine 1.14-fache Chance ($R^2 = .22$) festgestellt werden (Blossfeld et al., 2009; zitiert nach Hannover & Kessels, 2011). Jungen benötigten im Vergleich zu Mädchen somit im Durchschnitt bedeutsam mehr Kompetenzpunkte, um eine entsprechende Empfehlung für den Gymnasialübertritt zu erhalten. Der Ländervergleich zeigt allerdings, dass dieser globale Befund differenziert betrachtet werden muss: In Bremen und Bayern wurden nicht Jungen, sondern Mädchen trotz gleicher Leseleistung benachteiligt. Keine bzw. nur sehr geringe Unterschiede zwischen Mädchen und Jungen bei der Empfehlung für das Gymnasium machten Lehrer/-innen aus Sachsen, Rheinland-Pfalz, dem Saarland, Brandenburg und Niedersachsen. Nichtsdestotrotz existieren in 11 von 18 deutschen Bundesländern Unterschiede hinsichtlich der Gymnasialempfehlung zuungunsten von Jungen (Blossfeld et al., 2009).

Ergänzend zu diesen Befunden lässt sich aus den Ergebnissen der PISA-Studie 2006 (PISA-Konsortium Deutschland, 2007) entnehmen, wie viele Kompetenzpunkte Jungen und Mädchen

im PISA-Leistungstest erhalten haben und wie diese Kompetenzen in Noten umgesetzt wurden. Wie Hannover und Kessels (2011) darstellen, zeigten die Ergebnisse, dass Lehrkräfte Jungen im Fach Deutsch durchweg besser beurteilten, als Mädchen mit der gleichen Anzahl Kompetenzpunkte. So erhielten Mädchen z.B. mit 626 Kompetenzpunkten die Bestnote „1“, Jungen bekamen diese bereits ab 605 Kompetenzpunkten zugesprochen (*Anm.*: Richtung und Stärke der Notenvorteile für Jungen bleibt über die Notenstufen hinweg konstant). Ein konträres Bild zeigte sich in der Mathematik, wo Jungen wiederum mehr Kompetenzpunkte (615 Punkte) für die Bestnote erreichen mussten, als Mädchen (582 Punkte). Zur Erklärung dieser Befunde führen Hannover & Kessels (2011) an, dass Lehrkräfte offenbar bemüht sind, wahrgenommene Geschlechterunterschiede und geschlechtsbezogene Nachteile in Schulfächern entsprechend der gängigen Geschlechterstereotype auszugleichen und somit Jungen in einem stereotyp weiblichen Fach einen Kompetenzvorteil einzuräumen – nicht jedoch in einem Schulfach, in welchem das Stereotyp für die Leistungsfähigkeit von Jungen spricht. Eine generell schlechtere Platzierung von Jungen kann auf Grundlage dieser Befunde allerdings nicht abgeleitet werden. Einschränkend bleibt zusätzlich zu erwähnen, dass im Rahmen der PISA-Studie nur ein vergleichsweise geringer Ausschnitt an Schulfächern betrachtet wurde sowie dass Ergebnisse ausschließlich gemittelt über alle Schulformen vorliegen (was unter Einbezug des höheren Mädchenanteils auf Gymnasien zur Überlagerung von Effekten auf den einzelnen Schulformen führen kann). Darüber hinaus wiesen Voyer & Voyer (2014) in einer Metaanalyse (mit zumeist amerikanischen Stichproben) auf Notenvorteile für Mädchen hin, welche allerdings insbesondere in sprachlichen Domänen substantiell ausfielen. Eine genaue Beantwortung der Fragestellung, ob Jungen bei gleichen Kompetenzen im Vergleich zu Mädchen schlechter bewertet werden ist, demnach unter Einbezug der o.g. Befunde schwierig. Tendenziell sprechen diese aber für einen Vorteil von Mädchen zumindest beim Gymnasialübertritt und damit verbunden indirekt auch auf Ebene der Kompetenzen, da an höheren Schulformen auch mehr Kompetenzen erworben werden. Hannover & Kessels (2011) weisen jedoch darauf hin, dass bei der Interpretation dieser Befunde ebenfalls zu berücksichtigen ist, dass schulische Kompetenzen nicht mit Noten gleichgesetzt werden dürfen – für die Ermittlung einer Schulnote spielen häufig auch überfachliche Kompetenzen (Lern- und Arbeitsverhalten, soziale Kompetenzen) eine bedeutsame Rolle. Es ist somit notwendig, auch diese Faktoren entsprechend in die Diskussion einzubeziehen, da die Möglichkeit besteht, dass trotz ähnlicher kognitiver Kompetenzen zwischen den Geschlechtern aufgrund unterschiedlicher überfachlicher Kompetenzen differente Umsetzungen von Kompetenzen in Noten bestehen. Bereits an dieser Stelle soll darauf hingewiesen werden, dass aktuelle Befunde darauf hindeuten, dass überfachliche Kompetenzen eine zentrale Rolle bei der Betrachtung von Notenvorteilen zugunsten von Mädchen unter Kontrolle der kognitiven Fähigkeiten einnehmen (Duckworth & Seligman, 2006; Steinmayr & Spinath, 2008; Willingham & Cole, 1997). Im folgenden Abschnitt sollen – wiederum der Darstellung bei Hannover & Kessels (2011) folgend – die überfachlichen Kompe-

tenzen – welche auch ein Bestandteil von Schulnoten sind – genauer als weitere Erklärungsmöglichkeit für die Geschlechterdisparitäten in der Schule betrachtet werden.

Geschlechterunterschiede in überfachlichen Kompetenzen. Als eine weitere Erklärungsmöglichkeit für Geschlechterunterschiede bei der schulischen Leistung und dem Zertifikatserwerb werden von Hannover & Kessels (2011) Geschlechterdifferenzen in überfachlichen, sozialen Kompetenzen diskutiert, welche wiederum zu einer besseren Benotung führen und deshalb mit einem wahrscheinlicherem Besuch einer höheren Schulform einhergehen.

Kunter & Stanat (2002) untersuchten mehrbenenanalytisch mit der Stichprobe aus PISA 2000, welche Rolle das Geschlecht bei der Vorhersage zahlreicher Facetten sozialer Kompetenz – hier operationalisiert auf Ebene der Kooperation und Kommunikation – innehat. Als abhängige Variablen wurden u.a. Verantwortungsübernahme, Einhaltung arbeitsbezogener Normen, Empathie und Unterstützung von Mitschüler/-innen betrachtet. Es konnte gezeigt werden, dass das Geschlecht der Schülerinnen und Schüler der wichtigste Prädiktor für die Ausprägungen von z.B. aggressiven Orientierungen war, wobei Mädchen deutlich weniger aggressive Tendenzen berichteten als Jungen. Ähnliche Geschlechterunterschiede vergleichbarer Größenordnung waren bei fast allen untersuchten Aspekten von Kooperation und Kommunikation feststellbar. Auch unter Hinzunahme weiterer individueller Variablen, wie sozioökonomischer Status oder Migrationshintergrund, sowie Variablen auf Schulebene (Schulklima, Schulform) blieb Geschlecht der zentralste Prädiktor für die Vorhersage sozialer Verhaltensweisen im Schulkontext, insgesamt zeigten Mädchen durchweg positiveres Sozialverhalten als Jungen.

Weitere Hinweise auf Differenzen beim Sozialverhalten in Abhängigkeit des Geschlechts von Schülerinnen und Schülern lieferten Duckworth & Seligman (2006). Es wurde die Fragestellung aufgeworfen und analysiert, warum Mädchen durchweg bessere Noten in den Hauptfächern erwerben als Jungen und trotzdem Leistungstests oder Intelligenzmessverfahren keine Geschlechtsunterschiede feststellen können. Es wird vermutet, dass diese Diskrepanz durch eine höhere Selbstdisziplin und stärker ausgeprägtes selbstreguliertes Lernen bei Mädchen erklärt werden kann, da der Vorteil von Mädchen auf diesen Variablen für die Bewertung / Benotung zumindest ebenso relevant ist, wie akademische Leistungen. Bezüglich der Schulnoten erreichten Mädchen in dieser Untersuchung durchweg bessere Werte ($d = .48$ bis $.70$). Weiterhin wurde ersichtlich, dass bei nahezu allen eingesetzten Messverfahren zur Akquise von selbstreguliertem Lernverhalten Mädchen deutlich bessere Werte erzielten als Jungen – unabhängig davon, ob die Bewertung als Selbstbeschreibung, Lehrer- oder Elternurteil vorlag ($d = .32$ bis $.78$). Eine Mediationsanalyse bestätigte, dass der Zusammenhang von Geschlecht und Schulnoten durch die Selbstdisziplin vermittelt wurde. In einem ersten Regressionsmodell konnte aufgezeigt werden, dass der direkte Einfluss von Geschlecht auf Schulnoten hoch signifikant ausfiel ($B = 0.44$, $SE B =$

0.11, $r = .31$), jedoch unter Einbezug der Selbstdisziplin zurückging und Selbstdisziplin nunmehr die zentral bedeutsame Komponente darstellte ($B = 0.46$, $SE B = 0.07$, $r = .47$). In einer darauf aufbauenden Studie wurden alle o.g. Befunde repliziert. Zusätzlich konnte gezeigt werden, dass Jungen in einem Intelligenztest (OLSAT, Otis & Lennon, 1997) besser abschnitten als Mädchen. Eine ergänzende Mediationsanalyse unter Einbezug von Geschlecht und Kompetenzen (Testergebnis aus einem Leistungstest) baute auf die o.g. Analyse auf und replizierte den Befund, dass der signifikante Einfluss des Geschlechts auf die Schulnoten verschwindet, wenn zusätzlich in einem zweiten Schritt die Selbstdisziplin als Prädiktor aufgenommen wurde, wobei die Bedeutung der Kompetenzen erhalten blieb. Insgesamt liefert diese Studie bedeutsame Hinweise darauf, dass Jungen weniger Selbstdisziplin aufweisen als Mädchen und sich dies ungünstig auswirken kann, da neben fachlichen Kompetenzen auch weitere, überfachliche Kriterien für schulische Bewertungen herangezogen werden. Ergebnisse aus den Studien von Downey & Vogt Yuan (2005) unterstützen diesen Befund: Hier wurde festgestellt, dass in der Hauptsache unangepasstes Verhalten und mangelnde Disziplin in der Klasse den Befund aufklärten, warum Jungen schlechtere Bewertungen erhielten.

In einer längsschnittlichen Untersuchung von Ditton (2007) wurde die Bedeutung von personalen, sozialen und institutionellen Bedingungen auf den Kompetenzerwerb und den Schulübertritt am Ende der Grundschule untersucht. Hinsichtlich der Merkmale und Einstellungen der Schüler/-innen wurde festgestellt, dass die Lernfreude, also wie gern die Kinder zur Schule gehen und für die Schule arbeiten, vom ersten Messzeitpunkt (Ende Klasse 3) zum zweiten Messzeitpunkt (Ende Klasse 4) für beide Geschlechter bedeutsam abnahm. Der Vergleich von Jungen und Mädchen zeigte nichtsdestotrotz zu beiden Messungen eine deutlich geringere Lernfreude bei Jungen. Ein ähnliches Bild zeigte sich bei der Anstrengungsmotivation: Diese ging zwar ebenso von der ersten zur zweiten Messung bei beiden Geschlechtern in ungefähr gleichem Ausmaß zurück, jedoch waren Mädchen zu beiden Messzeitpunkten deutlich höher motiviert als Jungen (Kaufmann, 2007). Auf den Variablen extrinsische Leistungsmotivation und externe Attribuierungsmuster fanden sich hingegen keinerlei Geschlechterdifferenzen. Bezüglich des Zusammenhangs der o.g. Variablen konnte gefunden werden, dass Anstrengungsmotivation zwar signifikant negativ, dennoch mit nur geringfügigen Korrelationskoeffizienten mit Schulnoten (Deutsch, Mathe, Sachkunde) und positiv mit der Variable Übertrittsempfehlung auf ein Gymnasium korrelierte.

Ergänzend fanden sich in PISA-Daten des Jahres 2009 (Klieme et al., 2010) Hinweise auf Unterschiede in der Anwendung von Lernstrategien: Sowohl Wiederholungs- als auch Kontrollstrategien wurden in den meisten OECD-Ländern, so auch in Deutschland, von Jungen deutlich weniger häufig eingesetzt, als von Mädchen. Äquivalente Befunde existieren für das Lernstrategienwissen (Artelt, Naumann & Schneider, 2010). Weitere Studien lieferten ergänzende Erklä-

rungsfaktoren, wie eine höhere Motivation zur Arbeitsvermeidung bei Jungen (Steinmayr & Spinath, 2008) oder das männliche Schulkulturen weniger lernorientiert sind, als weibliche (Van Houtte, 2004). Ergänzend stellten Anders, McElvany & Baumert (2010) aus Perspektive von Lehrkräften fest, dass Mädchen einen (moderaten) Vorteil bei sozialen Fähigkeiten und Sozialverhalten sowie Motivation und Lerntugenden aufweisen, was wiederum deutlich negativ mit Schulnoten in Deutsch und Mathematik zusammenhing und mit einer höher wahrscheinlichen Übergangsempfehlung zum Gymnasium einher ging.

Insofern liefern die o.g. Arbeiten, die zum größten Teil der Darstellung von Hannover & Kessels (2011) folgend berichtet wurden, deutliche Hinweise darauf, dass Jungen zu ihrem Nachteil weniger soziale und motivationale Verhaltensweisen und Eigenschaften aufzeigen, die mit schulischen Anforderungen kompatibel sind und somit Geschlechterunterschiede bei überfachlichen Kompetenzen wichtige Erklärungsansätze für Geschlechterdisparitäten im Bildungswesen liefern können.

Schulisches Engagement als Bedrohung der männlichen Identität. In ihrem Überblicksaufsatz berichten Hannover & Kessels (2011) als einen weiteren möglichen Erklärungsansatz über ihre eigene Forschung. Zu dieser gehört auch das DFG-Forschungsprojekt „Sind Jungen die neuen Bildungsverlierer?“, in dessen Rahmen diese Dissertationsschrift entstanden ist und in dem ein weiterer Erklärungsansatz für das vergleichsweise schlechtere Abschneiden von Jungen diskutiert und einer empirischen Überprüfung zugeführt wird.

Wenn Schule als ein Ort begriffen wird, der nicht nur dem Erwerb von Fähigkeiten, Fertigkeiten (bzw. Kompetenzen) und Zertifikaten dient, sondern auch einen Rahmen bietet, in welchem Kinder und Jugendliche einen großen Teil ihres Alltagslebens verbringen (Krebs, 2008), ist zu vermuten, dass dieser auch bei der Herausbildung der eigenen Identität eine wichtige Rolle spielt (Kessels & Hannover, 2004, 2006). Dies betrifft dann natürlich auch diejenigen Bereiche der Identität, welche sich auf das eigene Geschlecht, oder genauer die eigene Femininität oder Maskulinität, beziehen. Vor diesem Hintergrund postulieren Hannover & Kessels (2011), dass eine zunehmende Schulabstinenz der Jungen dadurch initiiert und aufrechterhalten werden kann, dass schulisches Engagement insgesamt eher als weiblich wahrgenommen und daher durch Jungen vermieden wird, da es potentiell bedrohlich für die eigene Identität ist.

Skelton & Francis (2011) untersuchten in diesem Zusammenhang in einer qualitativen Studie mit begabten Jugendlichen, welche Jungen das Fach Englisch (ein eher weiblich konnotiertes Fach) als Lieblingsfach auswählten. Es konnte gefunden werden, dass bei ihren *Peers* sehr beliebte Jungen wesentlich häufiger eine weibliche Domäne als Lieblingsfach beschrieben als Jungen mit „nur“ guten Leistungen. Es scheint unter Einbezug dieses Befundes zentral, dass Jungen eine gesicherte Identität aufweisen und sich ihrer eigenen Männlichkeit durch eine hohe

Beliebtheit und soziale Eingebundenheit sicher sein müssen, damit von der Wahl einer feminin konnotierten Domäne keine soziale Bedrohung im Sinne einer Abwertung der eigenen Männlichkeit ausgeht. In einer Studie von Rentzsch, Schütz & Schröder-Abé (2011) zeigte sich, dass auf Vignetten beschriebene Jungen, die in der Schule Anstrengung / Engagement zeigten von den Probanden/-innen negativer bewertet wurden als Mädchen mit überdurchschnittlichen Lernergebnissen, was ebenso dafür spricht, dass Schule insgesamt von den Lernenden als weiblich konnotiert wahrgenommen wird.

Aufbauend auf diesen Befunden ist die zentrale Annahme im oben genannten DFG-Forschungsprojekt, dass die Herausbildung von schulischen Interessen (Fächerpräferenzen, Interesse an Schule allgemein, etc.) und auch von schulischem Engagement ganz allgemein (in welchem Maße werden schulische Angebote zur Kompetenzentwicklung genutzt) nachhaltig mit der Identitätsentwicklung im Jugendalter zusammenhängt (Kessels & Hannover, 2004, 2006). Jugendliche nutzen demnach die Angebote im Kontext Schule, um bestimmte Aspekte der eigenen Identität zu entwickeln und zu demonstrieren. Die Nutzung dieser Angebote hängt nicht ausschließlich mit dem Erwerb von fachlichen und überfachlichen Kompetenzen zusammen, sondern führt vielmehr auch zur Aneignung der mit schulischen Aktivitäten verbundenen sozialen Bedeutungen. Diese wirken sich auf das eigene Selbst aus, führen zu einer differenziellen Interessenentwicklung und münden schlussendlich kreislaufartig wieder in eine unterschiedliche Nutzung schulischer Angebote.

Mit diesem Erklärungsansatz wurde in der bisherigen Forschung vor allem erklärt, warum Mädchen sich bedeutsam seltener als Jungen den mathematisch-naturwissenschaftlichen Bereichen zuwenden (Hannover & Kessels, 2002; Kessels, 2002, 2012; Kessels & Hannover, 2006). Durch die stark männliche Konnotation des Schulfaches Physik (Kessels, 2002) sehen Mädchen sich in ihrer Selbstdefinition als Mädchen bedroht und stabilisieren ihre Geschlechtsidentität durch Vermeidung dieses Faches. Angewendet auf das Jungenproblem nehmen Hannover & Kessels (2011) an, dass auch Jungen die Schule als einen Ort nutzen, an dem sie ihre Männlichkeit durch ganz bestimmte, als männlich geltende Verhaltensweisen inszenieren. Dies führt in einem weiteren Schritt dazu, dass als weiblich geltende Verhaltensweisen wie strebsames Lern- und Arbeitsverhalten (schulisches Engagement) vermieden werden, um die männliche Geschlechtsidentität kommunizieren und stabilisieren zu können. Es wird angenommen, dass Jungen die Entwicklung und Inszenierung ihrer Männlichkeit darüber vornehmen, dass sie schulisches Engagement und auch typisch schulisches Sozialverhalten vermeiden, weil diese Aspekte bedrohlich auf die eigene Geschlechtsidentität wirken und die Selbstdefinition als „richtiger Junge“ gefährden. Durch die aktive Vermeidung derartiger Verhaltensweisen wird wiederum eine Stabilisierung der Geschlechtsidentität erreicht, welche allerdings – wie eingangs beschrieben – folgenreiche Konsequenzen für Jungen hat (Hannover & Kessels, 2011). Im Ausland werden dies-

bezügliche Ansätze bereits seit längerem diskutiert: „*laddish culture*“ (Gill, 2003), u.a. mit sehr maskulin besetzter Interessenverteilung, Abwertung von Weiblichkeit/Sexismus, Fokussierung auf als männlich geltende Attribute wie körperliche Stärke und Sportlichkeit und Rebellion gegen Autoritäten (u.a. Jackson, 2002; Warrington, Younger & Williams, 2000) werden hier mit der Herausbildung einer lernabträglichen Kultur und „*doing masculinity*“ assoziiert. Diese Aspekte und die daraus resultierenden Eigenschaften und Verhaltensweisen sind schlussendlich kaum mit schulischem Engagement und Erfolg vereinbar (Frosh et al., 2003; Jackson, 2002).

Zusammenfassend unterstützen die vorliegenden Befunde, dass für Jungen schulisches Engagement potentiell bedrohlich für die eigene Geschlechtsidentität wirken kann. Im Zuge der Frage nach Erklärungsmöglichkeiten für Geschlechterunterschiede in Kompetenzen, Noten und beim Zertifikatserwerb scheint die Annahme von Hannover & Kessels (2011) plausibel, dass Jungen die schulischen Angebote zur Inszenierung und Absicherung der eigenen Geschlechtsidentität verwenden bzw. gerade durch die Ablehnung von Schule und schulischem Engagement die eigene Identität als „richtiger Junge“ gleichsam dargestellt und abgesichert werden kann. Dies hat zur Folge, dass (wie im vorhergehenden Absatz beschrieben) eine Distanz zur Schule entsteht, einhergehend mit stärker unangepasstem Lern- und Arbeitsverhalten, das wiederum durch Lehrkräfte wahrgenommen wird und sich in entsprechenden Bewertungen niederschlägt, was sich im letzten Schritt in den eingangs dargelegten Befundmustern von schlechteren Bewertungen und geringer wertigem Zertifikatserwerb manifestiert. Somit liefert dieser Ansatz zusätzliche Erklärungsmöglichkeiten für Geschlechterdisparitäten im Bildungswesen über die bisher beschriebenen Ansätze hinaus und kann diese (insbesondere Befunde zu schlechterer Benotung und geringeren überfachlichen Kompetenzen) in einen größeren Gesamtzusammenhang integrieren.

2. Stereotype Threat-Theorie: Stereotypenbedrohung für Jungen i. d. Schule?

Wie der öffentliche Diskurs über "Jungen als Bildungsverlierer" Leistungsfähigkeit und Motivation bei Schüler/-innen beeinflussen kann.

Wenn sich Jungen in der Schule mit der Wahrnehmung von generell besseren Schulleistungen von Mädchen konfrontiert sehen, kann dies für sie zu einer Bedrohungssituation heranwachsen. Diese geht damit einher, dass aufgrund des wahrgenommenen Leistungsvorsprungs von Mädchen für die soziale Gruppe der Jungen die Befürchtung besteht, man könne dieses Vorurteil durch eigene Leistung und eigenes Verhalten bestätigen (*stereotype threat*; Steele & Aronson, 1995; Steele, 1997).

Die Theorie der Stereotypenbedrohung (*stereotype threat*) geht davon aus, dass Mitglieder einer (stigmatisierten) sozialen Gruppe mit dem Dilemma umgehen müssen, rein aufgrund ihrer Gruppenzugehörigkeit benachteiligt behandelt oder stereotyp beurteilt zu werden *und* Verhalten zu zeigen, welches die stereotypen Annahmen der Außengruppe bestätigt (Steele & Aronson, 1995). Insbesondere in Leistungssituationen kann diese Befürchtung leistungsmindernd wirken, nämlich genau dann, wenn diese Angst im Sinne einer sich selbst erfüllenden Prophezeiung wirkt und damit einhergehend Verhalten und Leistung im Sinne einer Bestätigung des Stereotyps beeinflusst (Steele, 1997). Die eigene Befürchtung, aufgrund der sozialen Gruppenzugehörigkeit kategorisiert und bewertet zu werden *und* diese durch eigenes Verhalten zu unterstützen wird somit zur Realität. Generell kann die Bedrohung durch Stereotype und damit verbundene Effekte jede soziale Gruppe betreffen (Stroessner & Good, 2011), jedoch weisen Aronson, Fried & Good (2002) mit der Einführung des Begriffs „*stereotype vulnerability*“ zusätzlich darauf hin, dass verschiedene Faktoren beeinflussen, wie schnell und vor allem wie stark ein Individuum auf die Bedrohung von Stereotypen reagiert (u.a. Identifikation mit der Gruppe, Kontrollüberzeugung, Wissen über Stereotype, persönliche Relevanz des von der Bedrohung betroffenen Kompetenzbereichs usw.).

Trotz der Vielzahl an potentiellen Möglichkeiten für betroffene Gruppen und daraus resultierenden Effekten von *stereotype threat* konzentrierte sich die bisherige Forschung (wie im folgenden Absatz ausführlicher gezeigt wird) hauptsächlich auf die Faktoren ethnische Herkunft und Geschlecht, genauer um Benachteiligung von afroamerikanischen Schüler/-innen und Studierenden in den USA, um wahrgenommene Leistungsvorteile von asiatischen Schüler/-innen sowie um Geschlechterdiskrepanzen im Fach Mathematik und damit verbundenen Nachteilen von Mädchen und Frauen. Zwar sind die Untersuchungen zum Zusammenhang von *stereotype threat* und Leistung in der Überzahl, dennoch existieren auch Befunde zur Beeinflussung anderer Faktoren wie Erhöhung kontraproduktiven Lernverhaltens (u.a. Reduktion der Lernzeiten in Anlehnung an Stone, 2002), *self-handicapping* (Steele & Aronson, 1995), Distanzierung von der

stereotypisierten Gruppe (Cohen & Garcia, 2008) und Vermeidung von potentiell bedrohlichen Situationen (Major, Spencer, Schmader, Wolfe & Crocker, 1998), Selbst-Blamage bei Fehlern (Koch, Müller & Sieverding, 2008) sowie Abwertung von Bedeutung und Validität von Leistungsaufgaben (Lesko & Corpus, 2006).

In einem für die *stereotype threat*-Forschung wegweisenden Experiment zeigen Steele und Aronson (1995), dass Stereotypenbedrohung die intellektuelle Leistungsfähigkeit negativ beeinflussen kann. Afroamerikanische Studierende zeigen unter der Experimentalbedingung (der vorgelegte, standardisierte Leistungstest wurde als diagnostisch für Intelligenz dargestellt) eine deutlich geringere Leistung als Afroamerikaner der Kontrollgruppe (in welcher der Test als von Intelligenz unabhängige Problemlöseaufgabe instruiert wurde) und europäisch-amerikanische Studierende beider Gruppen. Mit diesem Ergebnis wird ein Grundstein für die Betrachtung der negativen Auswirkungen von Bedrohung durch Stereotypen in *test taking situations* gelegt. Aufbauend auf dieser Arbeit entstanden eine Vielzahl ähnlicher Experimente, welche äquivalente Befunde zum Einfluss von Stereotypenbedrohung auf die Performanz leistungsbezogen-stereotypisierter sozialer Gruppen in verschiedenen Domänen nachweisen (Stroessner & Good, 2011) und damit die Robustheit der *stereotype-threat theory* unterstützen.

Tabelle 1 (aufgrund ihrer Größe im Anhang aufgeführt) soll ein Überblick über bisher veröffentlichte Arbeiten zur Stereotypenbedrohung, deren Fragestellung und zentraler Ergebnisse gegeben werden. Die Zusammenstellung der Befunde basiert auf den Meta-Analysen von Nguyen & Ryan (2008) sowie Picho, Rodriguez & Finnie (2013). Aufgenommen in diese Übersicht sind ausschließlich Untersuchungen, die einen klaren Bezug zum ursprünglichen Experiment von Steele und Aronson (1995) aufweisen, d.h. welche im Rahmen eines Leistungstests auf Bedrohungseffekte bei einer stereotypisierten Gruppe prüfen. Aus der Zusammenstellung der Befunde wird insgesamt deutlich, dass in der Hauptsache folgenden Fragestellungen nachgegangen wurde:

- Beeinflussung der Leistung von *Frauen*, insbesondere durch das Stereotyp von Nachteilen von Frauen in MINT-Fächern/Studiengängen: 115 (Teil-)Studien¹
- Leistungsdifferenzen zwischen *ethnischer Minorität* und Majorität, insbesondere in Bezug auf wahrgenommene Leistungsnachteile von ethnischen Minderheiten in den Vereinigten Staaten: 49 (Teil-)Studien¹

¹ Die im Anhang aufgeführte Tabelle 1 listet ausschließlich *veröffentlichte* Forschungsarbeiten auf. Hierbei ergeben sich geringere Zahlen publizierter Studien, und zwar 81 von 116 (Teil-)Studien zu Frauen in MINT sowie 25 von 116 (Teil-)Studien zum Einfluss ethnischer Herkunft auf standardisierte Testleistungen. Die im Text o.g. Studienzahlen ergeben sich durch Auszählung aller (auch unveröffentlichter) Arbeiten, welche im Rahmen von Metaanalysen herangezogen wurden.

Hinsichtlich der Ergebnisse lässt sich zu den einzelnen Fragestellungen zusammenfassend festhalten:

Geschlecht & Leistung: Gemittelt über alle Studien ergibt sich ein mittlerer Effekt des Leistungsabfalls in einem standardisierten Tests bei Frauen nach der Aktivierung von entsprechenden Geschlechterstereotypen mit im Mittel $d = |.36|$ (Nguyen & Ryan, 2008) bzw. $d = |.24|$ (Picho et al., 2013). Die Stärke des Effekts hing zusätzlich von der eingesetzten Aktivierungsstrategie (siehe Abs. 2.1.) ab, wobei Experimente mit subtiler Stereotypeninduktion die größten Effekte produziert haben. In der Richtung lässt sich durchweg eine nachlassende Leistung von Frauen im mathematisch-naturwissenschaftlichen Bereich unter *stereotype threat* feststellen, wohingegen die Effektstärken deutlich divergieren und von nicht vorhandenen ($d = |.01|$) bis zu sehr großen Effekten ($d = |1.7|$) variieren. An dieser Stelle sei angemerkt, dass keine in die Metaanalysen einbezogene Studie den Einfluss von Stereotypenbedrohung auf Männer in typisch weiblichen Leistungsdomänen untersuchte – zum aktuellen Zeitpunkt untersuchten dies: Hartley & Sutton (2013); Eckert & Imhof (2013); Hirnstein, Freund & Hausmann (2012); Keller (2007), Seibt & Förster (2004); Van Loo, Boucher, Rydell & Rydell (2013) und finden z.T. erste Hinweise auf Effekte einer Stereotypenbedrohung bei Männern in typisch weiblichen Leistungsdomänen.

Nur wenige Studien – welche nicht Bestandteil der o.g. Metaanalysen waren – fokussierten auf die Auswirkung einer Stereotypenbedrohung bei Jungen mit divergierenden Ergebnissen (u.a. Hartley & Sutton, 2013; Eckert & Imhoff, 2013), weiterhin entfallen wiederum einige Studien auf non-kognitive Bereiche, z.B. soziale Sensitivität (Koenig & Eagly, 2005), soziale Intelligenz (Cadinu, Maass, Lombardo & Frigerio, 2006) oder die Beurteilung des affektiven Gehalts von Wörtern (Leyens, Desert, Croizet & Darcis, 2000). Ferner existiert eine Studie von Weaver, Vandello & Bosson (2013) welche finanzielle Entscheidungen von Männern unter Stereotypenbedrohung untersuchte.

Ethnische Minoritäten: Es zeigt sich hier ein leicht stärkerer Effekt im Vergleich zu den Geschlechtsunterschieden mit $d = |.43|$ (Nguyen & Ryan, 2008), jedoch sprechen die Befunde auch hier für durchgehend leistungsmindernde Effekte von *stereotype threat* auf die standardisierte Testleistung von als benachteiligt stereotypisierten ethnischen Minoritäten. Ebenso variiert die Stärke des Effekts erheblich je nach Studie von $d = |.04|$ bis $d = |1.7|$. Noch stärker als in den einbezogenen Untersuchungen zu Geschlecht und Leistung beeinflusst die Strategie der Stereotypenaktivierung die durchschnittliche Effektstärke, wobei hier moderat explizite Aktivierungsstrategien die stärksten Effekte produzierten.

2.1. Aktivierung von stereotype threat

Die Aktivierung von *stereotype threat* und die damit verbundene, leistungsbeeinträchtigende Wirkung ist durch eine Vielzahl von Mechanismen möglich. In der aktuellen Forschungsliteratur wird in der Hauptsache nach der Explizitheit der Aktivierung differenziert. Hier bilden sich in den bisher durchgeführten Experimenten drei Gruppen der Stereotypenaktivierungen heraus. Bei einem ersten Mechanismus, der indirekten oder subtilen Aktivierung, wird lediglich der Kontext oder die Gruppenzugehörigkeit manipuliert. Ein ausdrücklicher Hinweis auf die fiktive oder tatsächliche Benachteiligung einer bestimmten Gruppe wird nicht gegeben. Vielmehr werden auf der Gruppenzugehörigkeit basierende Stereotypen und damit verbundene, angenommene Leistungsvor- oder -nachteile durch automatische oder unbewusste Mechanismen aktiviert (z.B. *Priming* der Geschlechtsidentität oder ethnischen Gruppenzugehörigkeit). Eine moderat explizite Aktivierung von Stereotypenbedrohung wird über die Vorlage von Stimulusmaterial umgesetzt, aus welchem bereits hervorgeht, dass eine bestimmte Subgruppe in einer Testsituation anders abschneidet, als eine Vergleichsgruppe. Es ist aber zusätzlich noch eine Interpretationsleistung zu erbringen, um die Richtung und ggf. die Stärke des Effekts feststellen zu können. Die eigentliche Bedrohungssituation wird somit erst durch eine bewusste Interpretation und Auslegung von gezeigtem Stimulusmaterial salient. Beispielweise wird eine derartige Aktivierungsstrategie in Arbeiten eingesetzt, in welchen ein Leistungstest zwischen zwei Gruppen deutlich differenziert und diese Differenzen ausgeführt oder dargestellt werden. Erst durch die eigenständige Interpretation der Unterschiede und die Aktivierung der eigenen Gruppenidentität erfolgt eine Bedrohung. Drittens existiert eine vollständig explizite Aktivierungsstrategie, mittels welcher direkt mitgeteilt wird, dass eine bestimmte Gruppe in einem Leistungstest schlechter abschneidet, als eine *Outgroup*. Bei dieser Strategie ist somit keine Interpretationsleistung mehr erforderlich, die Leistungsvorteile einer Gruppe werden deutlich kommuniziert. Dieser Ansatz referenziert auf den ursprünglichen Experimenten von Steele & Aronson (1995) und führt vor der Applikation einer Leistungsmessung deutlich aus, dass eine bestimmte Gruppe einen unbestrittenen Leistungsvorteil besitzt – die Bedrohung wird über einen vollständig bewussten Prozess induziert. Ein Beispiel für die Verwendung einer expliziten Strategie ist die deutliche Kenntlichmachung, dass in einem Leistungstest zur Messung mathematischer Fähigkeiten Frauen typischerweise wesentlich schlechter abschneiden, als Männer.

In zahlreichen Studien wurden alle beschriebenen Aktivierungsmechanismen eingesetzt und erweisen sich als wirksam zur Induktion einer Bedrohungssituation, wenngleich die Metaanalyse von Nguyen und Ryan (2008) darauf hindeutet, dass die Art der Aktivierung wiederum Auswirkungen auf die Stärke der gefundenen Effekte hat. Pauschal lässt sich allerdings nicht ableiten, welcher Aktivierungsmechanismus am besten geeignet scheint, um eine Bedrohung zu induzieren, da zahlreiche Randbedingungen ebenfalls eine nicht zu vernachlässigende

Rolle spielen. Hierzu zählen u.a., wie stark das zu aktivierende Stereotyp bereits internalisiert wurde, wie die wahrgenommenen Leistungsunterschiede attribuiert werden und wie stark eine Person von der Richtigkeit eines Stereotyps überzeugt ist – eine genauere Erläuterung der Einflussfaktoren auf die Wirkung von Stereotypenbedrohung findet im folgenden Abschnitt statt.

2.2. Mechanismen von stereotype threat

Trotz intensiver Forschungsbemühungen existieren derzeit nur Hinweise darauf, wie *stereotype threat* wirkt und wie die leistungshemmende Wirkung dieses Paradigmas vermittelt wird. Eine reduzierte Arbeitsgedächtnisleistung wird durch Beilock, Rydell & McConnell, 2007 postuliert. Negative Wirkungen von Stereotypenbedrohung treten in den Experimenten dieser Forschergruppe ausschließlich bei schwierigen Aufgaben auf und scheinen durch Verminderung der Arbeitsgedächtniskapazität vermittelt zu werden, wobei insbesondere phonologische Aspekte durch *stereotype threat* beeinträchtigt werden (vgl. auch: Régner, Smeding, Gimmig, Thinus-Blanc, Monteil & Huguet, 2010; Schmader & Johns, 2003). Weiterhin werden Ressourcenverbrauch durch Unterdrückung auftretender Emotionen und Kognitionen, wie z.B. Unterdrückung von Angst (Johns, Inzlicht & Schmader, 2008) oder eingeengte Wahrnehmung (Cadinu, Maass, Rosabianca & Kiesner, 2005) sowie psychologische Erregung und Anspannung (Ben-Zeev, Fein & Inzlicht, 2005) häufig im Zusammenhang mit leistungshemmenden Effekten durch Stereotypenbedrohung diskutiert.

In diesem Zusammenhang wiesen Spencer, Steele & Quinn (1999) darauf hin, dass der emotionalen Komponente Angst eine zentrale Rolle bei der Vermittlung von Effekten bei Bedrohung durch Stereotype zukommt. Es wurde bereits in früheren Studien der Autoren postuliert, dass die leistungshemmende Wirkung von *stereotype threat* dadurch entsteht, dass ein Gefühl der Ängstlichkeit als emotionale Reaktion auf die Bedrohung einen hemmenden Einfluss auf die Leistungsfähigkeit haben kann. Empirisch stellten Spencer et al. (1999) fest, dass Frauen in einem schwierigen Mathematiktest unter Herausstellung von geschlechtsspezifischen Ergebnissen in diesem Test auf der situationalen Komponente des *State-Trait Anxiety Inventory* (STAI, Spielberger, Gorsuch & Lushene, 1970) höhere Werte bezüglich der aktuellen Ängstlichkeit berichteten und zudem auch tatsächlich schlechtere Leistungen im Test erzielten. Als Erklärung für den leistungshemmenden Effekt in der bedrohten Gruppe wurden insbesondere die Beschäftigung mit der Angst, das negative Stereotyp zu bestätigen und ggf. angewendete *Coping*-Strategien diskutiert, welche die kognitiven Kapazitäten zur Aufgabenbearbeitung reduzieren. Zu berücksichtigen bleibt, dass andere Studien (z.B. Schmader, 2002) keine derartigen Effekte für Angst nachweisen konnten.

Eine eingeeengte Wahrnehmung sowie eine vermehrte Beschäftigung mit negativen Gedanken wiesen Cadinu et al. (2005) nach. In dieser Untersuchung wurden Frauen vor der Bearbeitung eines Mathematiktests gebeten, ihre Gedankengänge zu protokollieren. Es konnte gezeigt werden, dass in der Experimentalbedingung (explizite Aktivierung: Herausstellung, dass Männer in diesem Test deutlich besser abschneiden) Frauen wesentlich mehr negative Gedanken formulierten und diese einen hohen Bezug zur Testsituation aufwiesen, wie z.B. wahrgenommene eigene Unfähigkeit und schlechtere Mathematikkenntnisse. In der Kontrollgruppe (Hinweis, dass der Test keinerlei Geschlechtsunterschiede produziert) konnte dieses Ergebnismuster nicht aufgefunden werden. Im Ergebnis schnitten Frauen der Experimentalbedingung tatsächlich deutlich schlechter ab, als Probandinnen der Kontrollgruppe. Die Autoren/-innen resümierten, dass es somit gelungen ist, den Zusammenhang zwischen negativen Gedanken und schlechterer Testleistung nachzuweisen. Unklar bleibt allerdings, inwiefern dieses Ergebnis im Sinne einer *self-fulfilling prophecy* (Merton, 1948) interpretiert werden kann.

In weiteren Untersuchungen wurden ferner ergänzende Variablen zur Erklärung von *stereotype threat*-Effekten herangezogen. Es konnte nachgewiesen werden, dass in einer explizit bedrohten Experimentalgruppe von Frauen durch positives Feedback in einem Pretest keine verbesserte Leistungserwartung im Posttest verursacht wurde. Nur in einer nicht bedrohten Kontrollgruppe wurde durch positives Feedback in einer initialen Aufgabe auch eine entsprechend positive Leistungserwartung für folgende Aufgaben induziert. Es scheint daher, dass *stereotype threat* die Ausprägung einer positiven Leistungserwartung hemmt (Stangor, Carr & Kiang, 1998). In einer anderen Leistungsdomäne, dem Sport, konnte festgestellt werden, dass die subtile Aktivierung eines Stereotyps (Afroamerikaner sind naturgegeben sportlicher, als weiße Amerikaner) die Anstrengungsbereitschaft bedrohter Personen im Training signifikant reduzierte (Stone, 2002). Unklar bleibt bei der Interpretation dieses Befunds, ob diese Reduktion als „Rechtfertigung“ (im Sinne einer selbstwertdienlichen Attributionsmöglichkeit) unterschiedlicher, sportlicher Leistung dienen soll oder ob Ressourcen in einem scheinbar nicht zu gewinnenden Wettbewerb eingespart werden sollen.

Weitere Befunde zu Wirkungsmechanismen von *stereotype threat* stellt die Neuropsychologie bereit. In verschiedenen Studien konnte die Wirkung einer Stereotypenbedrohung als Stressor nachgewiesen werden, z.B. durch die Erhöhung der Herzfrequenz (Croizet, Després, Gauzins, Huguet, Leyens & Méot, 2004), des Blutdrucks (Blascovich, Spencer, Quinn & Steele, 2001) und des Cortisolspiegels (Townsend, Major, Gangi & Mendes, 2011) als biologische Indikatoren von Stress und Angst. Untersuchungen mit funktionaler Magnetresonanztomographie (fMRT) deuten auf eine erhöhte Aktivität des anterioren cingulären Cortex hin. Diese Bereiche sind für verschiedene Prozesse bei der Informationsverarbeitung des Menschen verantwortlich, so z.B. für die Verarbeitung sozialer und emotionaler Informationen und der Steuerung von Aufmerksamkeitsprozessen. Unter Stereotypenbedrohung wurde neben der Hyperaktivität die-

ses Hirnbereichs nachgewiesen, dass das Ausmaß der erhöhten, neurologischen Aktivität als signifikanter Prädiktor für die Performanz in einem anschließenden Leistungstest herangezogen werden konnte (Wraga, Helt, Jacobs & Sullivan, 2007) und dieses Areal somit eine wichtige Funktion für die neuropsychologische Vermittlung von *stereotype threat*-Effekten einzunehmen scheint.

Um eine Integration einer Vielzahl der soeben genannten Befunde bemühen sich Schmader, Johns & Forbes (2008) in ihrem „*Integrated Process Model of Stereotype Threat Effects*“. Da in den bisher durchgeführten Einzeluntersuchungen zum Thema *stereotype threat* nur wenige erklärende Variablen für die Wirkmechanismen mit erhoben wurden leiten die Autoren die unerwünschte Konsequenz einer Darstellung ab, dass lediglich eine Variable die Effekte von Stereotypenbedrohung auf die Leistung vermittelt. Das Modell bemüht sich daher um eine Zusammenführung von Befunden auf motivationaler, affektiver, physiologischer und kognitiver Ebene. In einem ersten Schritt wirkt sich die Bedrohung auf drei unterschiedliche, aber miteinander verbundene Mechanismen aus: i) eine physiologische Stressreaktion (Pfad *b*), ii) die Unterdrückung negativer Gedanken und Emotionen (Pfade *f* und *i*) sowie iii) aktive Monitoringprozesse (Pfad *d*), um weitere mögliche (bedrohende) Reize zu detektieren (siehe Abbildung 1). Nach dem Modell von Schmader et al. (2008) wirkt sich der Bedrohungsreiz nicht nur auf jede dieser drei genannten Komponenten aus, sondern es findet eine Interaktion aller Komponenten untereinander statt, sodass sich alle Teilsysteme auch gegenseitig beeinflussen (vgl. Doppelpfeile im Modell). Als zentrale, kognitive Komponente des Modells wird das Arbeitsgedächtnis angenommen, da dieses eine wichtige Rolle sowohl bei zahlreichen kognitiven Aufgaben als auch in sozialen Situationen einnimmt. Unter *stereotype threat* kann sich das Arbeitsgedächtnis aber entsprechend der Annahmen aus dem Modell nicht allein mit der Bearbeitung der Leistungsaufgabe beschäftigen, sondern ist auch aktiv an der Verarbeitung bedrohungsrelevanter Informationen beteiligt und nimmt zusätzlich eine abschirmende Funktion gegen diese Informationen ein. Trotz der Bemühungen des Arbeitsgedächtnisses um Schutz vor der Bedrohungssituation werden so kognitive Ressourcen gebunden, die nun für die Bearbeitung der eigentlichen Aufgabe nicht mehr zur Verfügung stehen, was entsprechend der Modellannahmen zu einer Verringerung der Leistung bei kognitiven und sozialen Aufgabenstellungen führt.

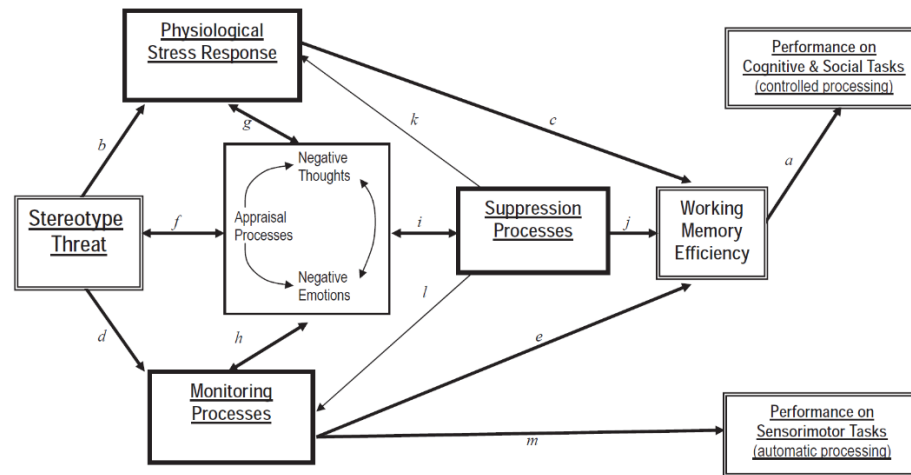


Abbildung 1: Integriertes Prozessmodell der Auswirkungen von *stereotype threat* auf die Leistung. Aus: Schmader, Johns & Forbes (2008), S. 337.

Ergänzend postuliert das Modell ebenfalls einen Pfad, wie automatisch verarbeitete Prozesse (sensorimotorische Aufgaben) beeinträchtigt werden: Durch die verstärkten (exekutiven) Monitoringprozesse unter *stereotype threat* werden nicht nur arbeitsgedächtnisrelevante Ressourcen für kognitive und soziale Aufgaben gebunden, sondern es gehen auf direktem Weg (Pfad *m*) hierdurch auch maßgebliche Anteile für derartige Aufgaben verloren.

Zusammenfassend lässt sich festhalten, dass das Modell von Schmader et al. (2008) neben der Integration zahlreicher, in experimentellen Arbeiten als vermittelnde Mechanismen identifizierter Komponenten dennoch einen Schwerpunkt auf die kognitiven Auswirkungen von *stereotype threat* setzt (Pfad *a*). Insbesondere Einflüsse der Bedrohung auf die Arbeitsgedächtnisfunktion durch die Verarbeitung und Abschirmung negativer Gedanken sowie Umgang mit Angst und Erregung sowie zusätzliche Einflüsse auf die Arbeitsgedächtnisfunktion durch vermehrten Einsatz von Monitoringstrategien und physiologische Stressreaktionen werden diesbezüglich angeführt. Wenngleich für jede im Modell enthaltene Komponente empirische Belege für die vermittelnde Wirkung und teilweise auch für das Zusammenspiel mehrerer Variablen (z.B. Regulation von Angst und beeinträchtigte exekutive Arbeitsgedächtnisfunktion; Johns et al., 2008) existieren, wurde das Gesamtmodell von Schmader et al. (2008) noch nicht pfadanalytisch überprüft. Ferner geht die Bedeutung motivationaler Variablen für die Vermittlung von Effekten nicht wie angekündigt (Schmader et al., 2008, S. 336) aus dem Modell hervor, sondern bleibt eher implizit und durch Hinzunahme weiterer Untersuchungen zu erschließen. Beispielsweise konnte Keller (2002) in einer Mediatoranalyse bestätigen, dass *self-handicapping* (entspricht im Modell von Schmader et al. (2008) Pfad *i*) den Zusammenhang zwischen Bedrohung und Testleistung vermittelte (zum Zusammenhang zwischen *self-handicapping* und Motivation vgl. Elliot & Church, 2003). Trotz der genannten Defizite bietet das Integrierte Prozessmodell erstmals einen Rahmen für die Einbettung verschiedener, in der Literatur diskutierter Variablen für die Vermittlung von *stereotype threat*-Effekten und eine Modellierung über das Zusammenspiel verschiede-

ner Komponenten. Durch den Einbezug zahlreicher, empirischer Befunde in die Modellierung ist es gelungen, den aktuellen Forschungsstand über die Wirkmechanismen der Stereotypenbedrohung in einen schlüssigen Gesamtzusammenhang zu integrieren.

Moderatoren von stereotype threat. Abschließend sei an dieser Stelle noch darauf verwiesen, dass neben Untersuchungen zu den soeben dargestellten, mediierenden Variablen von Effekten einer Stereotypenbedrohung auch verschiedene Arbeiten auf moderierende Einflussvariablen bei *stereotype threat* hinweisen, welche die Stärke der Effekte beeinflussen können.

Hinsichtlich der Aufgabenschwierigkeit weisen verschiedene Arbeiten darauf hin, dass die Angst ein negatives Stereotyp zu bestätigen bei leichten Aufgaben am ehesten zu einer Leistungsverbesserung führt, wohingegen Effekte durch Stereotypenbedrohung erst bei schwierigen Aufgaben auftreten (Steele, Spencer & Aronson, 2002; Keller, 2007). Es wird angenommen, dass bei leichten Aufgaben dem Modell von Schmader et al. (2008) folgend noch ausreichende Kapazitäten im Arbeitsgedächtnis vorhanden sind, um den Anforderungen leichter Aufgaben gerecht werden zu können. Dementsprechend wiesen z.B. Befunde von Steele et al. (2002) und Blascovich et al. (2001) keine Leistungsunterschiede auf, wenn es sich um hinreichend einfache Aufgaben handelt. Dem Prozessmodell von Schmader et al. (2008) folgend könnten der hohe Druck, das negative Stereotyp nicht zu bestätigen bei leichten Aufgaben sogar zu einem paradoxen Effekt einer Leistungssteigerung unter *stereotype threat* führen, wie sie z.B. bei O'Brien & Crandall (2003) aufgetreten sind: Bei einfachen Aufgaben erzielten Frauen in der Bedrohungsbedingung signifikant bessere Testleistungen als Probandinnen der Kontrollgruppe – das Ergebnismuster kehrte sich erst bei schwierigen Aufgaben um.

Als weiterer Moderator wird die Identifikation mit der Leistungsdomäne diskutiert. Befunden von u.a. Aronson, Lustina, Good, Keough, Steele & Brown (1999) sowie Martiny, Roth & Deaux (2011) folgend fielen Effekte einer Stereotypenbedrohung dann besonders substanziell aus, wenn eine Person sich hoch mit der bedrohten Leistungsdomäne identifizierte, d.h. wenn die Leistung im bedrohten Bereich einen wichtigen Teil des Selbstkonzeptes einer Person darstellte (zusammenfassend: Osborne & Jones, 2011). Kongruent fielen Effekte geringer bis gar gegenteilig aus, wenn sich eine Person nicht stark mit der bedrohten Leistungsdomäne identifizierte (Keller, 2007), was Cadinu, Maass, Frigerio, Impagliazzo & Latinotti (2003) ferner nicht nur für tatsächliche Leistungen, sondern auch Leistungserwartungen nachwies.

Die gleiche Wirkrichtung gilt ebenso für die Identifikation mit der eigenen *Ingroup*: aktuelle Untersuchungsergebnisse weisen diesbezüglich darauf hin, dass z.B. bei einer Bedrohung mit Geschlechterstereotypen die Ausprägung der Leistungseinbußen in einer Testsituation durch das individuelle Ausmaß der *gender identification* moderiert wird (Schmader, 2002; Wout, Danso, Jackson & Spencer, 2008). In beiden experimentellen Arbeiten konnte gezeigt werden, dass kongruent zur Identifikation mit der Leistungsdomäne ausschließlich dann leistungsmindernde

Effekte durch Stereotypenbedrohung auftraten, wenn sich eine bedrohte Person ausreichend hoch mit der eigenen Geschlechterzugehörigkeit identifizierte (z.B. „Ich bin stolz, Frau/Mann zu sein“, „ich identifiziere mich mit der Gruppe von Frauen/Männern“ usw.). Martiny, Roth, Jelenec, Steffens & Croizet (2012) demonstrierten ergänzend, dass moderierende Effekte der Gruppenidentifikation auch in vergleichsweise neu formierten Gruppen aufzufinden sind. Cadinu et al. (2003) kontrastieren die Befundlage und fanden in einer experimentellen Studie keinerlei Einfluss der *gender identification* auf Richtung und Stärke des *stereotype threat*-Effekts, wobei die Einteilung in *high* und *low identifiers* vergleichsweise undifferenziert über einen Median-Split vorgenommen wurde, während andere Analyseverfahren genauere Aussagen ermöglichen (vgl. Aiken & West, 1991). Bisher nicht beantwortet werden kann, welches Maß an Identifikation sowohl mit der Domäne als auch mit der Gruppe hinreichend ist, um die Induktion einer Bedrohungssituation durch Stereotypenaktivierung zu ermöglichen.

Zusätzlich existiert die Annahme, dass der Glaube an das Stereotyp, welches zur Bedrohung eingesetzt wird die Stärke der Effekte nachhaltig beeinflussen kann. Ausgehend von der Theorie von Steele & Aronson (1995) stellt eine feste Internalisierung des Stereotyps einhergehend mit einer eigenen Überzeugung, das Stereotyp sei richtig *keine* notwendige Bedingung für das Auslösen von *stereotype threat* dar. Der normative Druck, der durch das Stereotyp ausgelöst wird ist demnach nicht von der Akzeptanz des Individuums abhängig, sondern wirkt direkt als Bedrohung. Die aktuelle Befundlage ist hierzu allerdings als eher ambivalent zu betrachten und liefert Evidenz, dass der Glaube an ein Stereotyp durchaus einen Einfluss auf Effektstärken von *stereotype threat* haben kann: Während Schmader, Johns & Barquissau (2004) in einer experimentellen Studie feststellten, dass (bei salienter Geschlechtsidentität) die Leistungen von Testpersonen, die mehr an das zur Bedrohung eingesetzte Stereotyp glaubten, stärker beeinträchtigt wurde, wiesen hingegen Huguet und Regner (2009) Effekte von *stereotype threat* auch bei Probandinnen nach, die (zumindest deren Selbstauskünften nach) nicht an das negative Leistungsstereotyp glaubten. Somit kann nicht abschließend beantwortet werden, inwieweit von einer moderierenden Wirkung des Glaubens an ein negatives Stereotyp ausgegangen werden kann.

Insgesamt liegen – gemessen an den zahlreichen Befunden zu den Auswirkungen von *stereotype threat* – nur vergleichsweise wenige Arbeiten vor, welche Wirkmechanismen und Einflussfaktoren auf Bedrohungseffekte in der Datenakquise und -auswertung konsequent mit berücksichtigten. Ein verstärkter Einbezug von Moderator- und Mediatoranalysen in die *stereotype threat*-Forschung wäre daher wünschenswert, um bisher vorliegende Befunde zu Mediatoren und Moderatoren der Wirkung von Stereotypenbedrohung zu replizieren und ggf. zu ergänzen.

2.3. Verminderung von stereotype threat Effekten

Da die Auswirkung einer Stereotypenbedrohung keineswegs nur in einer kurzfristigen Beeinträchtigung der Testleistung besteht, sondern auch längerfristige Effekte wie die Vermeidung von potentiell bedrohlichen Leistungssituationen und damit einhergehend die nicht optimale Ausschöpfung eines Leistungspotentials auftreten (Major et al., 1998), rückte auch die Frage nach der Reduktion von Bedrohungseffekten in den Forschungsfokus.

In einem ersten Ansatz untersuchten Johns, Schmader & Martens (2005) die Auswirkungen einer umfassenden Information über das Phänomen und die Wirkungsweise von *stereotype threat*. In einem ersten Schritt wurde Männern und Frauen ein Leistungstest vorgelegt, welcher in einer Gruppe als Mathematiktest vs. als Problemlöseaufgabe vorgestellt wurde. Die Ergebnisse zeigten die typischen Auswirkungen einer Stereotypenbedrohung. In der Gruppe, in welcher der Leistungstest als Mathematikprüfung *gelabelt* wurde, schnitten Frauen im Durchschnitt deutlich schlechter ab, als Männer. Wurde der Leistungstest als Problemlöseaufgabe instruiert, verschwanden die Geschlechtereffekte und es zeigten sich vergleichbare Leistungen von Männern und Frauen. In einer aufbauenden Untersuchung wurden die Leistungsaufgaben erneut als Prüfung der mathematischen Kompetenz dargestellt, jedoch wurde umfassend darüber informiert, dass *stereotype threat* die Leistungsfähigkeit und das Testergebnis maßgeblich beeinflussen kann. Unter Hinzunahme dieser Informationen ergaben sich trotz des *Labelling* als Mathematiktest keine bedeutsamen Unterschiede zwischen den Geschlechtern mehr, was die Autoren dahingehend interpretierten, dass eine umfassende Information über die potentiellen Auswirkungen von Stereotypenbedrohungen negative Effekte (auf die Leistung) abschwächen oder gar auflösen kann.

McGlone & Aronson (2006) prüften, ob die zusätzliche Aktivierung von positiven Stereotypen zu einer Verminderung von Effekten der Stereotypenbedrohung führen kann. Dieser Ansatz basiert auf der Annahme, dass Individuen nicht nur exklusiv einer sozialen Gruppe zugehören, sondern multiple soziale Gruppenmitgliedschaften aufweisen. Demnach sollte getestet werden, ob die Aktivierung eines positiven Gruppenstereotyps die negative Stereotypisierung in einer Bedrohungssituation überlagern kann. Es konnte nachgewiesen werden, dass nach einem *Priming* auf die Geschlechtsidentität typische *stereotype threat*-Effekte im Sinne einer verminderten Leistung von Frauen in einem Test zur mentalen Rotation auftraten. Wurde jedoch vor dem Test darauf fokussiert, dass alle Teilnehmenden einer hochselektiven Privathochschule angehören konnten keine Leistungsunterschiede zwischen Männern und Frauen mehr gefunden werden. Es konnte damit demonstriert werden, dass negative Effekte durch *stereotype threat* vermindert werden können, wenn ein anderes, positives Gruppenstereotyp salient gemacht wird.

Durch die Betonung der Veränderbarkeit von kognitiven Fähigkeiten und Intelligenz erreichten Aronson et al. (2002) eine Verminderung von Effekten durch Stereotypenbedrohung. In einem Experiment mit afroamerikanischen Studierenden, über welche ein negatives Leistungstereotyp in der amerikanischen Gesellschaft existiert, wurden diese instruiert, in einer potentiell bedrohlichen Leistungssituationen Intelligenz und kognitive Fähigkeiten als ein individuelles und vor allem veränderliches, trainierbares Konstrukt wahrzunehmen. Diese gedankliche Strategie führte dazu, dass insbesondere die Gruppe afroamerikanischer Studierenden insgesamt mehr Freude bei Lernprozessen zeigte, ein größeres akademisches Engagement aufwies und auch bessere Studienleistungen im Sinne von besseren Noten erzielten, als Probanden/-innen zweier Kontrollgruppen. Hieraus kann geschlussfolgert werden, dass negative Effekte von *stereotype threat* dadurch vermindert werden können, indem die bedrohte Domäne (in diesem Falle: Intelligenz / kognitive Leistungsfähigkeit) nicht als „naturegegeben“ oder starr wahrgenommen wird, sondern als willentlich veränderbare Größe interpretiert werden sollte. Ferner bietet diese Form des Umgangs mit Stereotypenbedrohung auch im Falle von Misserfolgen die Möglichkeit einer lernförderlichen Attribution.

Die Wirksamkeit von *self-affirmation* (Steele, 1988) zur Abschwächung von Bedrohungseffekten wiesen u.a. Cohen, Garcia, Apfel & Master (2006) nach. Den Autoren/-innen zufolge soll durch Selbstbestätigung der bedrohte Selbstwert gestützt und damit verbunden negative Leistungseffekte abgeschwächt werden. In zwei experimentellen Studien mit subtiler Stereotypenaktivierung wurde überprüft, inwiefern die Leistung im schulischen Kontext (Noten) bei afroamerikanischen Schüler/-innen positiv durch *self-affirmation* beeinflusst werden kann. In einer Interventionsgruppe wurde den Schüler/-innen neben zu bearbeitenden Leistungsaufgaben zusätzlich zu einem bestimmten Zeitpunkt eine Liste mit verschiedenen Eigenschaften vorgelegt (z.B. gute Beziehungen zu Freunden haben). Aus dieser Liste war eine Eigenschaft auszuwählen, die als am wichtigsten erachtet wurde (in der Replikation: die drei wichtigsten Eigenschaften). Zusätzlich ist ein kurzer Absatz zu verfassen, warum das ausgewählte Item für die eigene Person von großer Bedeutung ist. Ziel der Intervention war es somit, die Integrität des Selbst trotz Stereotypenbedrohung aufrecht zu erhalten. Die Kontrollgruppe sollte aus der Liste die am wenigsten favorisierten Eigenschaften auswählen und in der schriftlichen Aufgabe darauf eingehen, warum diese Eigenschaft für andere Personen wichtig sein könnte. Die Ergebnisse zeigten, dass die durch Stereotype bedrohten, afroamerikanischen Schüler/-innen der Interventionsbedingung nach der Intervention bessere Noten aufwiesen, als Probanden/-innen der Kontrollbedingung (Treatmenteffekte: $B = 0.26 - 0.30$, entspricht einer Reduktion von 40% des *racial achievement gap*). Bei Unterteilung in Subgruppen nach (*pre-intervention*) Leistungsniveau zeigte sich, dass insbesondere Personen mit initial schwacher oder mittelmäßiger Leistung besonders von der Intervention profitieren konnten. In einem Wortvervollständigungstest wurde auf einer

anderen Ebene ebenfalls die Wirksamkeit der Intervention nachgewiesen: Schüler/-innen der Interventionsgruppe generierten weniger stereotypenbehaftete Wörter (z.B.: _ACE = FACE statt RACE) als Probanden/-innen der Kontrollbedingung. Dies indiziert zusätzlich die Verringerung der Verfügbarkeit von Stereotypen. Hervorzuheben bleibt die Tatsache, dass die Autoren/-innen mit einer kleinen und vergleichsweise unaufwändigen Interventionsstrategie beachtliche Effekte für die Verminderung negativer Effekte von *stereotype threat* erzielen konnten. Das Ergebnis bezüglich der leistungsverbessernden und *stereotype threat* hemmenden Wirkung einer derartigen Intervention wurde mit Studierenden der Medizin durch Woolf, McManus, Gill & Dacre (2009) repliziert.

Ergänzend konnten Walton & Cohen (2007) zeigen, dass das Zugehörigkeitsgefühl zu einer bestimmten Domäne eine zentrale Rolle für die Vulnerabilität gegenüber *stereotype threat* spielt. In einer Intervention, welche die Zweifel an der Zugehörigkeit zu einer (in diesem Experiment: akademischen) Domäne abmilderte und somit ein stärkeres Zugehörigkeitsgefühl zum akademischen Kontext etablierte wiesen Studierende ethnischer Minderheiten deutlich bessere Studienleistungen und mehr Erfolgszuversicht auf, als vor der entsprechenden Intervention:

“Simply normalizing doubts about social belonging – presenting them as common across racial groups – and portraying such doubts as temporary rather than stable made Black students’ sense of fit less dependent on the quality of their day, increased their engagement in achievement behavior (e.g., time spent studying) and, it seems, improved their GPA” (Walton & Cohen, 2007, S. 93).

Es konnte anhand einiger, exemplarischer Studien verdeutlicht werden, dass negative Auswirkungen von Stereotypenbedrohungen keineswegs unveränderliche Tatsachen darstellen und dass leistungsmindernde Effekte durch den Einsatz von diversen Interventionsstrategien effektiv gemindert werden können. Für die Beseitigung akuter Bedrohungssituationen, insbesondere im akademischen Kontext, scheinen die vorgeschlagenen Interventionen durchaus wirkungsvoll und ökonomisch einsetzbar. Inwiefern diese Maßnahmen jedoch geeignet sind, Bedrohungen durch Stereotype dauerhaft „auszuschalten“ oder einer raschen Extinktion unterliegen bleibt allerdings noch in längsschnittlichen Forschungsarbeiten zu überprüfen.

2.4. Stereotype reactance

Stereotypenbedrohungen gehen nicht notwendigerweise ausschließlich mit Leistungsverminderungen einher, sondern können sogar unter Umständen gegenteilige Effekte hervorrufen. Unter dem Begriff „*stereotype reactance*“ wird subsummiert, dass bedrohte Individuen psy-

chologische Reaktanz (Brehm, 1966) zeigen und damit einhergehend motiviert sind, dem Stereotyp entgegen zu arbeiten – also alle verfügbaren Ressourcen aufzuwenden, um stereotypeninkonsistentes Verhalten zu zeigen und die negativen Leistungserwartungen nicht zu bestätigen.

In vier experimentellen Studien untersuchten Kray, Thompson & Galinsky (2001) die Auswirkungen einer Stereotypenbedrohung sowie den Einfluss der Art der Stereotypenaktivierung auf die Verhandlungsleistung von Studierenden der Ökonomie und der Psychologie, insbesondere im Hinblick auf a) leistungsmindernde Effekte durch Stereotypenbedrohung und b) leistungssteigernde Effekte durch Reaktanz. In einem ersten Experiment konnte gezeigt werden, dass die Darstellung einer Aufgabe (welche typischerweise als eher männlich konnotiert wahrgenommen wird / Preisverhandlungen) als diagnostisch für Leistung dazu führt, dass Frauen stereotypenkonform schlechtere Leistungen erzielten. Die Aufgabe war ausschließlich in Mann-Frau-Dyaden zu bearbeiten. Die Ergebnisse zeigen, dass Männer 10% höhere Gewinne erwarteten (und diese auch erzielten) sowie extremere Verhandlungsstrategien aufwiesen, z.B. extreme Eröffnungsangebote aussprachen. Die Interaktion in den Verhandlungsdyaden wurde in einem zweiten Experiment genauer betrachtet, zusätzlich wurden Mann-Mann-Dyaden in das experimentelle Design aufgenommen. Ebenfalls ergänzend wurde randomisiert in der Experimentalgruppe einem Verhandlungspartner mitgeteilt, dass der Erfolg der Aufgabenbearbeitung u.a. von der eigenen Durchsetzungsfähigkeit und einem konsequenten Vertreten der eigenen Interessen (also eher männlich stereotypisierten Eigenschaften) abhängt. In den gleichgeschlechtlichen Dyaden konnten keine Unterschiede bezüglich der Verhandlungsleistung gefunden werden, wohingegen in zweigeschlechtlichen Dyaden Männer erneut bessere Leistungen erzielten als Frauen und die Größe dieses Effekts nach dem Hinweis über die (geschlechtsstereotypen) Einflussfaktoren nochmals stärker zugunsten der Männer ausfiel. Bei Frauen führte die zusätzliche Bedrohung durch die vermeintlichen, den Erfolg beeinflussenden (männlichen) Eigenschaften nicht zu einer weiteren Verminderung der Leistung, was die Autoren/-innen dahingehend interpretieren, dass die reine Beschreibung der stark männlich konnotierten Aufgabe bereits zur Induktion von *stereotype threat* geführt hat und durch die Hinzufügung der männlichen Verhaltenseigenschaften für die erfolgreiche Aufgabenbearbeitung keine maßgebliche Erhöhung der Bedrohungsintensität eintritt – womöglich weil derartige Eigenschaften durch die Probandinnen bereits antizipiert wurden. Aufbauend auf diese Vorarbeiten zum Nachweis, dass die durch die Autor/-innen konstruierte Aufgabe in der Lage ist, *stereotype threat*-Effekte auf verschiedenen Wegen auszulösen wurde in den darauf aufbauenden Experimenten 3 und 4 das Phänomen der *stereotype reactance* genauer untersucht. Hierzu wurde zusätzlich die Art der Stereotypenaktivierung manipuliert. In einer Bedingung fand eine eher implizite Aktivierung statt (Ausführung der o.g. Verhaltenseigenschaften für eine erfolgreiche Aufgabenbearbeitung) vs. eine explizite Bedingung, in welcher zusätzlich zu den Verhaltenseigenschaften klar konsta-

tiert wurde, dass Männer i.d.R. eine bessere Leistung erzielen. Es zeigte sich, dass in Mann-Frau-Dyaden in der expliziten Bedingung deutliche Effekte von Reaktanz bei Frauen auftraten, indem Frauen in dieser Bedingung durchweg bessere Leistungen erzielten, als Männer (höhere Eröffnungsangebote, effizientere Verhandlungsstrategien und bessere monetäre Erlöse). Hieraus wird ersichtlich, dass eine klare und starke Ausführung von Geschlechterunterschieden eine notwendige Bedingung für die Genese von *stereotype reactance* zu sein scheint. Experiment 4 replizierte die genannten Befunde zur Reaktanz (erneut bessere Leistung von Frauen in der expliziten Aktivierungsbedingung) und fokussierte verstärkt auf kognitive Mechanismen, die diesen Effekt vermitteln. Es konnte gezeigt werden, dass in Mann-Frau-Dyaden die eigene Emotionalität als potentiell hinderlicher Verhandlungsfaktor in der expliziten Bedingung bei Frauen wesentlich geringer ausfällt (31% Zustimmung), als bei Männern (58% Zustimmung). In der Kontrollgruppe kehrt sich dieses Muster um, hier stimmen 46% der Frauen dieser Frage im Vergleich zu nur 22% der Männer zu. Es scheint, als distanzieren sich Frauen vom klassischen Rollenbild (*disidentification*) und damit von potentiell erfolgsvermindernden Verhalten und bilden eine stereotypeninkonforme, aber leistungsförderliche Identität aus. Die vergleichsweise nachlassende Leistung von Männern wird damit begründet, dass diese sich unter expliziter Repräsentation von Leistungsvorteilen vermehrtem Druck ausgesetzt sehen und aus diesem Grund die Leistungsfähigkeit abnimmt: Durch die Befürchtung, ein positives Stereotyp nicht bestätigen zu können tritt gleichermaßen eine Bedrohungssituation ein, welche die Leistungsfähigkeit der Männer hemmt (*choking under pressure*; Baumeister & Showers, 1986). Es konnte anhand dieser Arbeiten gezeigt werden, dass subtile Aktivierungsmechanismen von *stereotype threat* die Leistungsfähigkeit einer bedrohten Subgruppe im Sinne der klassischen Experimente abschwächen können, aber explizitere Aktivierungsstrategien jedoch gegenteilige Effekte produzieren können und somit eine Leistungssteigerung einer eigentlich einer Bedrohung ausgesetzten, sozialen Gruppe nach sich ziehen kann. Das vorherrschende Bild einer generell leistungsmindernden Wirkung von Stereotypenbedrohung konnte damit um eine wichtige, weitere Wirkrichtung ergänzt werden.

In zwei weiteren Experimenten stellten aufbauend auf diese Befunde Kray, Reb, Galinsky & Thompson (2004) zusätzlich fest, dass in Verhandlungssituationen, in welchen die Verhandlungsmacht manipuliert wurde, nach expliziter Stereotypenaktivierung (verbunden mit postulierten Vorteilen für Männer) Probanden und Probandinnen in der Bedingung mit hoher Verhandlungsmacht effektiver, härter und erfolgreicher verhandelten. Das Ausbleiben von Geschlechtereffekten bewerten die Autoren/-innen als Zeichen für ein aktives Reagieren von Frauen gegen das induzierte Stereotyp. Das zweite Experiment ging mit dem Hinweis auf eine bessere Leistung von Frauen einher, erneut zeigen sich hinsichtlich der erworbenen Ergebnispunkte maximale Leistungen in der Bedingung mit expliziter Stereotypenaktivierung, jedoch ohne

Hinweis auf Geschlechterunterschiede hinsichtlich der Leistung. In diesem Falle scheinen Männer nach der expliziten Aktivierung von Stereotypen, die für eine bessere Leistung von Frauen sprechen *stereotype reactance* gezeigt zu haben. Zusammenfassend wird festgehalten, dass diese Ergebnisse die Hypothese unterstützen, dass eine explizite Aktivierung von Geschlechterstereotypen die Teilnehmenden dazu bewegt hat, als Reaktion im Sinne einer *stereotype reactance* eine Anpassung an das positive Stereotyp einhergehend mit einer Assimilation entsprechender Verhaltensweisen vorzunehmen.

Gupta, Turban & Bhawe (2008) finden ebenso, dass die explizite Darstellung des Arbeitsfeldes der Unternehmensführung als Männerdomäne bei Frauen zu Reaktanz i.S.v. gesteigerten Intentionen, in diesem Bereich tätig werden zu wollen mündet, wohingegen eine eher implizite Assoziation mit männlich konnotierten Verhaltenseigenschaften mit nachlassenden Intentionen von Frauen, in diesem Bereich arbeiten zu wollen einhergeht. Ohne stereotypische Informationen bestätigt sich wiederum, dass Männer deutlich höhere Intentionen aufweisen, als Frauen. Insgesamt spricht auch dieser Befund für das Vorhandensein von *stereotype reactance* in Abhängigkeit von der Expliztheit der Stereotypenaktivierung wie bei Kray et al. (2001, 2004) und erweitert deren Befund auf einen Kontext, in welchem kein leistungsbezogenes Maß als abhängige Variable untersucht wurde.

In einem weiteren experimentellen Setting fanden Oswald & Harvey (2000/2001) ebenfalls Hinweise auf *stereotype reactance*. In einem Experiment mit weiblichen Studierenden wurde überprüft, ob eine feindselige Umgebung die Mathematikleistung negativ beeinträchtigt. Hierzu wurde der Versuchsraum mit einem abwertenden Poster präpariert, auf welchem eine Frau an einem einfachen mathematischen Problem scheitert, während ein Mann ein komplexes Problem scheinbar mühelos löst. Befindlichkeitsmessungen bestätigten die frustrierende Wirkung dieser Manipulation. Zusätzlich wurde als Faktor variiert, ob während der Instruktion für den Mathematiktest darauf hingewiesen wird, dass Männer und Frauen gleich gute Leistungen in diesem Test zeigen oder nicht. Die Ergebnisse unterstützen die Annahmen von *stereotype reactance*: In der Bedingung, in welcher das hostile Poster gezeigt und *nicht* auf gleiche Leistungen von Männern und Frauen hingewiesen wurde, erzielten die untersuchten Frauen deutlich bessere Ergebnisse i.S.v. mehr bearbeiteten Aufgaben und mehr richtig gelösten Problemen im Vergleich zur Gruppe, welche ebenfalls dem Poster ausgesetzt war, jedoch auf die gleiche Testleistung von Männern und Frauen hingewiesen wurde. Offensichtlich wurden die Probandinnen der Experimentalgruppe motiviert, entgegengesetzt zu der durch das Poster vermittelten Aussage zu reagieren. In der Kontrollbedingung (ohne hostiles Poster) erzielten Frauen eine bessere Testleistung *mit* dem Hinweis, dass in diesem Test typischerweise keine Geschlechterunterschiede produziert werden.

In einer Untersuchung zu Auswirkungen von Stereotypenbedrohungen von Männern im verbalen Bereich setzten Hirnstein et al. (2012) einen Wortflüssigkeitstest ein und fanden ebenfalls Hinweise auf Reaktanz: In der *stereotype threat*-Bedingung wiesen Männer eine deutliche bessere Testleistung auf, als Männer der Kontrollbedingung. Die zum endgültigen Nachweis berechnete Varianzanalyse weist allerdings keine signifikante Interaktion Geschlecht x Bedingung aus, post-hoc Einzelvergleiche wurden nicht berichtet. Bei Betrachtung der Leistungsmittelwerte ist davon auszugehen, dass die fehlende Interaktion darauf zurückzuführen ist, dass Frauen in beiden Bedingungen nichtsdestotrotz Leistungsvorteile aufwiesen. Werden nur die Mittelwerte und Standardabweichungen von Männern der Experimental- und der Kontrollgruppe betrachtet, kann jedoch davon ausgegangen werden, dass sich die Leistungen signifikant i.S.v. besseren Leistungen der Männer in der Bedrohungsbedingung unterscheiden ($M_{EG} = -.45$, $SD = .12$; $M_{KG} = -.03$, $SD = .16$; $t(34) = -9.01$, $p < .001$, $d = -2.97$; eigene Berechnung, Mittelwerte und Standardabweichung wurden der Abbildung 1 aus Hirnstein et al., 2012 entnommen). Somit kann daher angenommen werden, dass Männer der Experimentalgruppe eine Reaktion im Rahmen von *stereotype reactance* zeigen. Durch die erstmalige Untersuchung von Männern als benachteiligte Gruppe geht aus diesem Befund ferner hervor, dass die bisher berichteten Befunde zu *stereotype reactance* keine geschlechtsspezifischen Phänomene sind, sondern gleichermaßen bei beiden Geschlechtern auftreten.

Abschließend weisen Befunde von Leyens et al. (2000) darauf hin, dass *stereotype reactance* nicht zwingenderweise mit einer Leistungsverbesserung einhergeht. In einer Worterkennungsaufgabe sollten die Probanden/-innen u.a. die Affektivität bestimmter Wörter beurteilen, d.h. schnellstmöglich angeben, ob einem Wort eine affektive Bedeutung beikommt. Eine Subgruppe erhielt zusätzlich zu einer allgemeinen Instruktion die Information, dass es ein gut gesicherter Fakt ist, dass Frauen affektbezogene Informationen besser und schneller verarbeiten können, als Männer. Männliche Versuchspersonen dieser Gruppe zeigten tatsächlich eine verminderte Leistung im Sinne einer gesteigerten Fehlerrate. Bei genauerer Betrachtung kann festgestellt werden, dass die Fehler bei der Einschätzung der Affektivität der Stimuluswörter nicht auf Auslassungen zurückzuführen sind (*error of omission*), sondern ausschließlich auf einer viel zu häufigen Bewertung von Wörtern als affektiv (*error of commission*) basieren. Männer der Experimentalgruppe haben also in der Bedrohungssituation insofern Reaktanz gezeigt, als dass sie die Reizschwelle für die Bewertung eines Wortes als affektiv deutlich abgesenkt haben, um in keinem Falle einen Reiz fälschlicherweise nicht als affektiv zu klassifizieren (*zero-miss-Strategie*). Die Autor/-innen interpretieren dieses Ergebnis als Versuch, dass der sozialen Gruppe anhaftende Stereotyp zu widerlegen und liefern somit einen wichtigen Hinweis auf die Tatsache, dass *stereotype reactance* zumindest im Kontext von Entscheidungsaufgaben eine erhöhte

Rate von Fehlentscheidungen – der Signalentdeckungstheorie folgend eine erhöhte Rate „falscher Alarme“ – nach sich ziehen kann.

2.5. Die andere Seite der Medaille: *Stereotype lift* und *stereotype boost*

In Abgrenzung zur o.g. Darstellung über Reaktanz bei Stereotypenbedrohung und zur weiterführenden Erklärung positiver Leistungseffekte sollen an dieser Stelle kurz die Konzepte *stereotype lift* (Walton & Cohen, 2003) und *stereotype boost* (Shih, Pittinsky & Ho, 2012) erläutert werden.

Beim Phänomen des *stereotype lift* erfolgt gleichsam zur *stereotype reactance* eine Leistungssteigerung, jedoch nicht in der eigentlich bedrohten Gruppe wie bei *stereotype reactance*, sondern in der (nicht bedrohten) *Outgroup*. In einer Metaanalyse von 42 experimentellen Studien zu Stereotypenbedrohung fanden Walton & Cohen eine Effektstärke von $d = .24$ in Richtung einer besseren Leistung der nicht bedrohten sozialen Gruppe, wenn Leistungsunterschiede (z.B. auf Basis von ethnischer Herkunft oder Geschlecht) postuliert wurden. In Kontrollgruppen, in welchen kein entsprechender Hinweis auf Leistungsunterschieden auf Grundlage der o.g. Faktoren gegeben wurde, divergierten die Leistungen hingegen nicht zwischen den untersuchten Gruppen. Hierbei genügte bereits ein subtiler Hinweisreiz i.S.v. Angaben zur hohen Diagnostizität eines Tests für kognitive Leistungen, um entsprechend leistungsverbessernde Wirkungen in der nicht stereotypisierten *Outgroup* nachzuweisen. Erst eine explizite Entkräftung von Unterschieden in der Testleistung bzw. die deutliche Herausstellung von Geschlecht oder Herkunft als irrelevant für die Testleistung führten dazu, dass die nicht bedrohte Gruppe keine Leistungsverbesserung aufwies.

Shih et al. (2012) bezeichnen ein Phänomen der Leistungssteigerung als *stereotype boost*, beziehen sich dabei allerdings auf positive Stereotype über die *Ingroup*, welche die leistungssteigernde Wirkung vermitteln. Somit kontrastieren die Arbeiten zu *stereotype boost* in der Hauptsache entsprechende Untersuchungen zu *stereotype threat* und weisen nach, dass die Aktivierung von positiven Stereotypen eine leistungssteigernde Wirkung haben kann. Den Befunden zur Stereotypenbedrohung folgend betonen Shih, Ambady, Richeson, Fujita & Gray (2002) ebenso die wichtige Rolle der Art, wie die positiven Stereotype hervorgerufen werden. So zeigen die Autoren/-innen, dass eine subtile Aktivierung zu den oben beschriebenen, leistungssteigernden Effekten führen. Eine explizite Aktivierung hingegen geht mit einer verminderten Leistung einher, was auf einen zu hohen Leistungsdruck durch die Stereotypeninduktion zurückgeführt wird (*choking under pressure*). Neben dem Hauptbefund entsprechender Studien zu *stereotype boost*, dass kongruent zum Leistungsabfall bei negativer Stereotypisierung der *Ingroup* auch umgekehrt ein leistungssteigernder Effekt bei positiver Stereotypisierung auftreten kann, wird

ebenso eine Erklärungsmöglichkeit für bessere Leistungen der nicht stereotypisierten Gruppe geboten: Shih et al. (2012) postulieren, dass mit der Induktion eines negativen Stereotyps über eine Außengruppe gleichsam positive Stereotype die *Ingroup* betreffend mit induziert werden. Die Theorie des *stereotype boost* geht also ferner von leistungssteigernden Effekten aus, wenn die nicht von einer negativen Stereotypisierung betroffenen Personen gleichsam ein positives Stereotyp für die Eigengruppe wahrnehmen und dementsprechend mit einer Leistungssteigerung reagieren. In Abgrenzung zu Walton & Cohen (2003) wird allerdings angenommen, dass die negative Stereotypisierung einer Außengruppe nicht direkt zu einer Leistungssteigerung führt, sondern zunächst unter Einsatz kognitiver Ressourcen in ein positives Stereotyp die *Ingroup* betreffend uminterpretiert werden muss.

Trotz divergierender Ansichten über die zugrundeliegenden Mechanismen liefern beide Ansätze, *stereotype lift* und *stereotype boost* wertvolle, ergänzende Erklärungen für in der bisherigen Forschungsliteratur berichtete, leistungsbezogene Effekte im Zusammenhang mit der Aktivierung von Stereotypen. Neben der negativ stereotypisierten Gruppe scheinen entsprechende Stereotypenaktivierungen auch die eigentlich nicht vom Stereotyp betroffene Gruppe (positiv) in ihrer Leistung zu beeinflussen (u.a. Walton & Cohen, 2003; Johnson, Barnard-Brak, Saxon & Johnson, 2012). Ferner konnte gezeigt werden, dass eine Umkehrung von *stereotype threat* möglich scheint und bei entsprechender Änderung der negativen Valenz eines Stereotyps für die *Outgroup* in eine positive Valenz für die *Ingroup* eine gesteigerte Leistung resultieren kann (u.a. Shih et al., 2002; Wraga, Duncan, Jacobs, Helt & Church, 2006; Levy, 1996). Da bei Stereotypisierung einer *Outgroup* beide Paradigmen zur Erklärung einer Leistungssteigerung der nicht stereotypisierten Gruppe herangezogen werden können, erfolgt die Verwendung beider Begriffe – trotz der o.g. konzeptionellen Unterschiede – teilweise noch uneinheitlich.

2.6. Anwendung auf die Problemstellung

Wie aus dem bisher gegebenen Überblick deutlich wurde, ist das Konzept und die Auswirkungen von *stereotype threat* von hoher Relevanz in verschiedenen Lehr-Lern-Settings. Insbesondere durch die nachgewiesenen, leistungshemmenden Effekte wird deutlich, dass die Konfrontation mit einer negativen Leistungserwartung auch im Lernkontext Schule im Hinblick auf Leistungsunterschiede eine wertvolle Erklärungsgrundlage bietet. Der bisherige Forschungsfokus konzentrierte sich dabei eher auf Nachteile von Frauen in naturwissenschaftlich-technischen Bereichen (Nguyen & Ryan, 2008). Getriggert durch die mediale Aufbereitung der „Jungenproblematik“ in der Schule (vgl. Abs. 1.1.) und das damit einhergehend gesteigerte Interesse am Bildungserfolg von Jungen seitens der Öffentlichkeit und in der Forschung (Weaver-Hightower, 2003) verlagert den Forschungsfokus zunehmend.

Eine zentrale Fragestellung ist, ob Jungen durch die öffentliche Darstellung ihres angeblichen Scheiterns und dem damit einhergehenden Wandel der Geschlechterverhältnisse negative Leistungsstereotypen internalisieren und diese wiederum in einer Leistungssituation als Bedrohung im Sinne von *stereotype threat* wahrgenommen werden. Überträgt man die bisherigen Befunde auf die Problemstellung der vorliegenden Arbeit, kann gemutmaßt werden dass bei Vorliegen entsprechender negativer Leistungsstereotypen über Jungen ebenso leistungsmindernde Effekte in einer Testsituation auftreten können, wie bei Frauen im MINT-Bereich. Inwieweit die generelle Bedrohung von Jungen als Schulversager *across the board* auch zu generell leistungsmindernden Effekten führt, soll im Rahmen dieser Arbeit geprüft werden. Erste Hinweise auf eine Übertragbarkeit bisheriger Ergebnisse zu *stereotype threat* auf die Jungenforschung liefern Hartley & Sutton (2013). In einer ersten Studie wurde untersucht, inwieweit Schüler/-innen die Ansicht vertreten, dass Mädchen bessere Schulleistungen erbringen. Hierzu wurden verschiedene Szenarien vorgelesen – mit Schildern, welche die Silhouette eines Jungen oder eines Mädchen zeigten wurde gemessen, welche Geschlechtergruppe für das vorgelesene Szenario die bessere Leistung erbringt. Es konnte gezeigt werden, dass bei Szenarien, die auf den schulischen Kontext bezogen waren signifikant mehr Mädchen nominiert wurden, als Jungen. Bei männlichen Versuchspersonen verstärkte sich dieser Effekt sogar über den Verlauf der Schuljahre. Darauf aufbauend wurde in einer zweiten Studie ein Testheft vorgelegt, welches aus standardisierten Aufgaben zum Lesen, Schreiben und zur Mathematik bestand. Dieses wurde in der Experimentalgruppe so eingeführt, dass Mädchen typischerweise eine bessere Leistung erbringen. In der Kontrollgruppe wurde keinerlei geschlechterbezogener Hinweis gegeben, sondern nur auf die individuelle Leistung fokussiert. Die Ergebnisse machten deutlich, dass in der Experimentalgruppe Jungen eine deutlich schlechtere Leistung in allen drei gemessenen Bereichen aufwiesen, als Mädchen und als Jungen der Kontrollgruppe. In der Kontrollgruppe divergierten die Leistungen nicht zwischen den Geschlechtern. Abschließend prüfte Studie 3, ob die Widerlegung des Stereotyps eine Leistungsverbesserung für Jungen zur Folge hat. Hierzu erhielt die Experimentalgruppe die Information, dass Mädchen und Jungen in diesem Test typischerweise die gleiche Leistung erbringen (Kontrollbedingung kongruent zu Studie 2). Die Ergebnisse verdeutlichen, dass Jungen der Experimentalbedingung im Vergleich zur Kontrollbedingung eine verbesserte Leistung zeigten, d.h. die Widerlegung des Stereotyps entfaltete einen leistungsförderlichen Effekt. Diese Studie untersuchte jedoch nur eine englische Stichprobe im Grundschulalter. Unklar bleibt, inwieweit sich dieser Effekt auch im Sekundarbereich fortsetzt und ob sich mit zunehmender Schulerfahrung bei älteren Schüler/-innen das Stereotyp vermehrt ausdifferenziert oder so generell wie beschrieben bestehen bleibt.

Eckert & Imhof (2013) zeigen hingegen einen gegenteiligen Befund zu *stereotype threat* und Schulleistungen von Jungen. Es wurden an einer reinen Jungenstichprobe drei experimentelle Bedingungen umgesetzt, impliziter *stereotype threat* (Test zur Messung verbaler Fähigkei-

ten im Fach Deutsch), explizite Bedrohung (Test, in welchem Mädchen typischerweise bessere Leistungen erbringen) und eine Kontrollbedingung (Negierung von Geschlechterunterschieden). Zusätzlich wurde die Identifikation mit der Domäne Lesen und die Identifikation mit dem eigenen Geschlecht gemessen. Danach wurde ein kontinuierlicher Lesetext vorgelegt und die Leistung mit darauf bezogenen Kontrollfragen erfasst. Es zeigte sich, dass weder ein Einfluss der Experimentalbedingungen, noch ein Einfluss der erhobenen Identifikation mit der Domäne und dem eigenen Geschlecht die Leseleistung der Jungen bedeutsam beeinflusst hat. Zusammenfassend wird resümiert, dass die schädlichen Auswirkungen von *stereotype threat* bei Mädchen nicht *per se* auf Jungen übertragbar sind. Diese Ergebnisse stehen somit im Kontrast zu den Annahmen aus der *stereotype threat*-Theorie. Problematisch hierbei ist allerdings zu betrachten, dass keine Kontrollgruppe (Mädchen) in die Studie einbezogen wurde und ein Großteil der Stichprobe (zwei Drittel) einem Gymnasium, also der höchsten deutschen Schulform, entstammten. Es bleibt fraglich, inwieweit diese Gruppe der (gemessen am besuchten Schultyp) doch recht erfolgreichen Jungen durch die dargestellten Stereotype ausreichend beeinflusst werden konnte. Ebenfalls war die Kontrollbedingung eigentlich eine *stereotype nullification*-Bedingung und damit keine „echte“ Kontrollbedingung – hier wäre es notwendig gewesen, keinerlei Angaben über Geschlechterunterschiede bei der Testleistung zu machen.

Trotz dieser ambivalenten Befunde wird deutlich, dass das Konzept der Stereotypenbedrohung als eine mögliche Erklärung bei der Analyse des Bildungserfolgs von Jungen zu berücksichtigen ist. Insbesondere im Hinblick auf die Öffentlichkeit der geführten Debatte um das Versagen von jungen in der Schule liefert *stereotype threat* eine Möglichkeit, Leistungsunterschiede besser verstehen zu können. Ferner sollen im Rahmen dieser Arbeit die oben dargestellten offengebliebenen Fragen der beschriebenen Studien zur Leistung von Jungen unter Stereotypenbedrohung bearbeitet werden, um diese Forschungslücke zu schließen.

Wenngleich – wie im vorausgegangenem Abschnitt deutlich wurde – das Phänomen *stereotype threat* in Bezug auf leistungsmindernde und leistungssteigernde Effekte in großer Bandbreite empirisch untersucht wurde, sind hingegen Effekte auf Hintergrundvariablen zum Lernen als Wissenserwerb wie z.B. Lern- und Leistungsmotivation bzw. motivationale Zielorientierungen vergleichsweise weniger untersucht. Der folgende Abschnitt soll kurz theoretische Annahmen zum Zusammenhang von Lernmotivation und Leistung darstellen und darauf aufbauend mögliche Einflüsse einer Stereotypenbedrohung auf derartige Variablen darstellen.

3. Jenseits von Leistung: Stereotypenbedrohung und Motivation

3.1. Theoretische Grundlagen

Neben Auswirkungen von Stereotypenbedrohung auf die Leistung in einer bedrohten Domäne beschäftigten sich weitere Forschungsarbeiten mit Auswirkungen von *stereotype threat* auch auf damit korrespondierende Variablen wie die Lern- und Leistungsmotivation. Auf Motivation als eine zentrale Wirkkomponente von Stereotypenbedrohung weisen Martiny & Götz (2011) mit ihrem Modell zur Wirkung von *stereotype threat* hin: Effekten auf Leistungsebene werden gleichrangig Auswirkungen auf die prospektive Motivation gegenübergestellt.

Schiefele (2008) führt als Formen von Lernmotivation vier Komponenten an: intrinsische und extrinsische Motivation, Interesse und Leistungsmotivation. Entsprechend den Grundannahmen von *stereotype threat* ist demnach zu erwarten, dass eine lernende Person durch eine bedrohliche Wirkung der Stereotypen befürchtet, durch eigenes Verhalten (d.h. insbesondere durch die eigene Leistung) ein negatives Stereotyp über die Eigengruppe zu bestätigen (Keller, 2007).

Dies würde implizieren, dass Lernverhalten verstärkt extrinsisch motiviert wird, d.h. das Erbringen einer Leistung wird von außen dadurch motiviert, die Leistungsfähigkeit der eigenen Gruppe zu demonstrieren was impliziert, dass die intrinsische Motivation unter *stereotype threat* eingeschränkt wird (d.h. das Erbringen der Leistung erfolgt nicht auf der Grundlage, dass durch den Prozess an sich Freude und Zufriedenheit im Individuum entsteht). Dadurch, dass durch ein negatives Stereotyp die Leistungsfähigkeit einer Person in einer bestimmten Domäne infrage gestellt wird, ist weiterhin zu erwarten, dass individuelles Interesse im bedrohten Leistungsbereich eher gering ausgeprägt ist: Prenzel, Krapp und Schiefele (1986) gehen in ihrer Person-Gegenstands-Theorie davon aus, dass Interesse begünstigt wird, wenn der Gegenstand der Lern-Leistungssituation eine positive, gefühlsbezogene Valenz besitzt, eine hohe wertbezogene Valenz vorliegt (entspricht einer hohen, subjektiven Bedeutung) und eine „Selbstintentionalität“, also eine vollkommene Abwesenheit äußerer Zwänge vorliegt. Alle postulierten Voraussetzungen für hohes Interesse, mit Ausnahme der subjektiven Bedeutsamkeit, werden durch eine Stereotypenbedrohung verletzt und sollten demnach in vermindertes Interesse münden.

Weiterhin liegt eine Auswirkung von *stereotype threat* auf die Leistungsmotivation und Zielorientierungen nahe. Dem Risiko-Wahl-Modell von Atkinson (1957) zufolge wird die Auswahl einer Leistungsaufgabe durch zwei übergeordnete Tendenzen bestimmt (diese bestehen wiederum aus 3 Unterkategorien, auf welche an dieser Stelle nicht weiter eingegangen werden soll), nämlich die Tendenz, Erfolg anzustreben und die Tendenz, Misserfolg zu vermeiden. Den Grundannahmen von *stereotype threat* folgend sollten durch Stereotype bedrohte Personen ein besonders ausgeprägtes Motiv nach Misserfolgsvermeidung aufweisen, um das negative Stereo-

typ keinesfalls zu bestätigen und daher vergleichsweise einfachere Aufgaben in Leistungssituationen präferieren. Ergänzend weist Dweck (1986) in ihrer Theorie der Leistungsmotivation darauf hin, dass Leistungsziele in zwei verschiedene Bereiche aufgeteilt werden können: *performance goals* stellen das Bedürfnis dar, im Vergleich mit einer Gruppe möglichst gut abzuschneiden, wohingegen *mastery* oder *learning goals* ausschließlich auf den Lernprozess und den daraus zu erwartenden Lerngewinn fokussieren – hier steht die Weiterentwicklung eigener Fähigkeiten im Vordergrund (Dweck, 1986; Dweck & Elliott, 1983). Aus diesem sozial-kognitiven Ansatz wurde abgeleitet, dass bei Wahrnehmung einer geringen Fähigkeit in einer bestimmten Fachdomäne (worauf das Paradigma von *stereotype threat* basiert) eher ein „*avoid challenge behavior pattern*“ zu erwarten ist (Dweck & Leggett, 1988).

Da die Theoreme von Atkinson (Bedürfnis nach Erfolg und Vermeidung von Misserfolg) hierbei nicht berücksichtigt werden, erweitert Elliot (1999) das Modell um Annäherung (Herstellung oder Erhaltung positiver Zustände) und Vermeidung (Ausweichen negativer Ereignisse). Tabelle 2 gibt einen Überblick über die resultierenden Zielorientierungen, wenngleich sich die Differenzierung von Lernzielen nach Annäherung und Vermeidung *nicht* durchgesetzt hat und daher von einem trichotomen Zielansatz ausgegangen wird (Spinath, 2009).

Tabelle 2: 2x2Leistungsziel-Modell nach Elliot & McGregor (2001)

	Lernziel	Leistungsziel
Annäherung	Aufgaben als Herausforderung / Lerngelegenheit (→ langfristiger Erfolg)	Zeigen eigenen Wissens und Könnens (→ kurzfristiger Erfolg)
Vermeidung	Erlerntes Wissen und Fähigkeiten beibehalten/absichern	Verbergen mangelnden Wissens und Könnens (→ kurz- und langfristige Leistungsprobleme)

Diesem Schema zufolge ist in bedrohlichen Situationen wie unter *stereotype threat* insbesondere ein Vermeidungs-Leistungsziel zu erwarten, aber (z.B. bei sehr einfachen Aufgaben) auch ein Annäherungs-Leistungsziel alternativ denkbar.

Urhahne (2008) postuliert sieben Themenfelder der Lernmotivation, namentlich (1) klassische Leistungsmotivationstheorie, (2) erweiterte Erwartungs-Wert-Theorien, (3) Attributionstheorien, (4) Zielorientierungskonzepte, (5) Theorien intrinsischer Motivation, (6) Volitionsmo-
delle und (7) Theorien sozialer Motivation und ordnet diese in ein Modell zur Darstellung motivierten Handelns ein. Dieses Modell konstruiert sich aus den Faktoren Person, Situation, Ziel,

Handlung, Ergebnis und Folgen. Im Folgenden soll skizziert werden, welche Auswirkungen von *stereotype threat* für jede Art von Lernmotivation abgeleitet werden können².

Bei der klassischen Leistungsmotivationstheorie kommt der darin postulierten Erfolgserwartung (Atkinson, 1957) eine besondere Bedeutung zu. Diese sollte unter Stereotypenbedrohung negativ ausfallen, sich nach dem Handlungsmodell auf die antizipierten Folgen auswirken (Antizipation von Misserfolg) und schlussendlich in einem Vermeidungs-Leistungsziel der Person münden.

Erwartungs-x-Wert-Modelle setzen einen erwarteten Wert einer Handlungsalternative in Verbindung mit einer Erwartung, die entsprechende Handlung auch positiv ausführen zu können. Die Handlung mit dem größten Erwartungs-Wert-Produkt sollte dann gewählt und ausgeführt werden (Rudolph, 2003). Unter *stereotype threat* wäre zu vermuten, dass die Einschätzung der Erfolgswahrscheinlichkeit (Person → Ergebnis) ungünstig ausfällt. Ebenso ist abzuleiten, dass Ergebniserwartung (Handlung → Ergebnis) negativ konnotiert wird, d.h. es wird angenommen, dass eine Handlung mit einer bestimmten Wahrscheinlichkeit ein ungünstiges Ergebnis liefern wird (z.B. das Leistungsergebnis entspricht stereotyper Erwartung). Dies wirkt sich schlussendlich auch auf die Ergebnis-Folge-Erwartung aus, die Antizipation von Misserfolg führt zur Antizipation negativer Folgen (z.B. Bestätigung des Stereotyps). Diese Mechanismen können in der Gesamtheit betrachtet erklären, wie Stereotypenbedrohung eine Handlung und damit einhergehend die Auseinandersetzung mit einem stereotypisierten Lerngegenstand hochgradig unwahrscheinlich machen.

In Bezug auf Attributionstheorien als Einflussgröße auf die Lernmotivation wird festgehalten, dass die Zuweisung internal-stabiler Gründe (z.B. Fähigkeit) bei Erfolg und die Zurückführung von Misserfolg auf externale Faktoren (z.B. mangelnde Anstrengung) besonders günstige Auswirkungen haben (Weiner, 1986). Da unter *stereotype threat* ein negatives Leistungsstereotyp aktiviert wird, welches zumeist auf internalisierten Fähigkeitsunterschieden und damit auf internal-stabilen Gründen basiert (z.B. „Frauen haben eine geringere Fähigkeit in Mathematik“) kann vermutet werden, dass unter Stereotypenbedrohung vermehrt ungünstigere Attributionsschemata als Folge eines Leistungsergebnisses auftreten, was sich negativ auf Selbstwert und Lernmotivation der Person auswirkt.

Konzepte zur Volition gehen davon aus, dass eine hohe Motivation nicht hinreichend für die Erreichung eines (Lern-) Ziels ist, sondern dass volitionale Prozesse notwendig sind, um einen Handlungsvorsatz zu bilden und Handlungsschwierigkeiten angemessen zu begegnen (Kuhl, 1987). Es kann vermutet werden, dass *stereotype threat* ggf. volitionale Prozesse und damit ver-

²Anmerkung: Auswirkungen von *stereotype threat* unter Berücksichtigung von Zielorientierungskonzepten und Theorien intrinsischer Motivation wurden bereits auf Seite 47f. erläutert und werden daher nicht nochmals aufgegriffen.

bundene Prozesse der Regulation von Kognition, Emotion und Verhalten negativ beeinflussen können (Inzlicht, Tullett & Gutsell, 2011; Major & O'Brien, 2005).

Soziale Lernmotivationstheorien betonen die Wichtigkeit sozialer Eingebundenheit in eine Lerngemeinschaft und sozialer Unterstützung für die Lernmotivation. Obwohl mit der Akzentuierung auf soziale Eingebundenheit gleichsam wie bei Stereotypenbedrohung ein Gruppengefühl angesprochen wird, scheint es nichtsdestotrotz schwierig, aus der Übertragung dieser theoretischen Richtung in das Handlungsmodell von Urhahne (2008) Ableitungen für die Beeinflussung sozialer Lernmotivation durch *stereotype threat* zu treffen, weshalb auf die Generierung von Annahmen an dieser Stelle verzichtet werden soll.

Zusammenfassend lässt sich aus den Konzepten zu Lern- und Leistungsmotivation zumindest theoretisch ableiten, dass Stereotypenbedrohung auch umfassende Auswirkungen auf motivationale Prozesse in Lern- und Leistungskontexten hat. Im Folgenden soll dargestellt werden, auf welche Art und Weise eine empirische Überprüfung dieser Annahmen stattgefunden hat, wobei ein Schwerpunkt empirischer Evidenz auf Befunden zum Konzept der Beeinflussung von Leistungsmotivation und Zielorientierungen liegt.

3.2. Empirische Befunde

Vick, Seery, Blascovich & Weisbuch (2008) untersuchten biopsychologische Variablen als Indikatoren der Auswirkung einer Stereotypenbedrohung und liefern erste Hinweise auf Einflüsse von *stereotype threat* auf motivationale Lagen („*challenge*“ vs. „*threatened*“ motivational state). Wenn ein Individuum eine Aufgabe als mit den vorhandenen Ressourcen bewältigbar wahrnimmt, resultiert nach Vick et al. (2008) eine herausfordernde motivationale Orientierung. Eine Wahrnehmung, dass eine vorliegende Aufgabe nicht mit den eigenen, vorhandenen Ressourcen zu bewältigen ist, resultiert den Autoren/-innen nach in einer bedrohten motivationalen Lage. Um diese Annahme zu überprüfen wurde ein schwieriger Mathematiktest durchgeführt, welcher angeblich Geschlechterunterschiede zuungunsten von Frauen gefunden hat (*threat*) vs. keine Geschlechterunterschiede produziert (*no threat*). Als Indikatoren für eine herausfordernde vs. bedrohliche motivationale Lage wurden einem biopsychosozialen Modell (Blascovich & Tomaka, 1996) folgend über den kardialen Output (z.B. Herzrate) und ein Maß der Kontraktion vs. Dilatation des Gefäßsystems (*TPR, total peripheral resistance*) während der Aufgabenbearbeitung operationalisiert. Es konnte gezeigt werden, dass Probandinnen in der *threat*-Bedingung entsprechend der Annahmen des biopsychosozialen Modells (Blascovich & Tomaka, 1996) ein kardiovaskuläres Befundmuster zeigten, welches auf eine bedrohte motivationale Lage hinweist, wohingegen Männer ein entgegengesetztes Befundmuster einhergehend mit einer herausfordernden motivationalen Lage zeigten. In der Kontrollbedingung kehrte sich dieses Befundmuster um, da Männer hier durch die implizite Negation von Leistungsvorteilen in eine

bedrohte motivationale Lage versetzt wurden, wohingegen Frauen unter der Antizipation von Geschlechtergleichheit in den Testleistungen einen kardiovaskulären Befund zeigten, welcher auf eine herausfordernde motivationale Lage hindeutet. Diese Studie präsentiert erstmals Ergebnisse zu Auswirkungen einer Stereotypenbedrohung auf die Motivation *während* der Testbearbeitung. Der genaue Wirkmechanismus auf kognitiver oder motivationaler Ebene bleibt durch die Befunde zwar zunächst unklar, ebenso wie der tatsächliche Effekt auf die Leistung, da keinerlei Leistungsdaten in die Analysen einbezogen wurden. Nichtsdestotrotz wird es durch die Autoren/-innen unter Rückbezug auf die Daten als plausibel erachtet, dass eine bedrohte motivationale Lage die Vulnerabilität gegenüber einer verminderten Leistung erhöht und somit motivationale Konsequenzen nach sich zieht.

Die Auswirkung von *stereotype threat* auf die intrinsische Motivation wurde häufig nur im Zusammenhang mit Vermeidungs-Leistungszielen (s.u.) untersucht (Thoman, Smith, Brown, Chase & Lee, 2013). So konnte gezeigt werden, dass eine motivationale Lage, die der einer Stereotypenbedrohung entspricht, d.h. deren zentrales Ziel es ist einen potentiellen Misserfolg zu vermeiden, dazu führt, dass die intrinsische Motivation untergraben wird (Elliot & Harackiewicz, 1996). Eine Metanalyse von Rawsthorne & Elliot (1999) liefert weitere unterstützende Evidenz für die Annahme, dass die Verfolgung von Leistungszielen einen nachteiligen Effekt auf die intrinsische Motivation hat. Eine direkte Auswirkung einer Stereotypenbedrohung auf die intrinsische Motivation liegt nach einer Recherche in der Datenbank PsycINFO derzeit nicht vor.

Smith, Sansone & White (2007) liefern Befunde zum Einfluss von Stereotypenbedrohung auf das Interesse. Weibliche Studierende, welche zum Zeitpunkt der Untersuchung noch kein akademisches Hauptfach gewählt haben, wurden gebeten einen „Test zur Beurteilung von Computer-Fähigkeiten“ zu bearbeiten, welcher in der Experimentalbedingung als Test dargestellt wurde, der Geschlechterunterschiede produziert vs. in der Kontrollbedingung Geschlechterunterschiede negiert. Zufällig wurde den Studierenden ein Vermeidungs-Leistungsziel (schlechtes Abschneiden vermeiden), ein Annäherungs-Leistungsziel (positives Testergebnis anstreben) oder keine Leistungszielorientierung induziert. Hernach wurden den Studierenden einige aufgabenrelevante Grundlagen erläutert, um eine gemeinsame Wissensbasis sicherzustellen. Nach diesem einführenden Tutorium waren einige Aufgaben aus dem Bereich Computertechnik und Programmierung zu bearbeiten. Abschließend wurde das Interesse erfragt, mehr zu den soeben präsentierten Inhalten zu lernen. In der Experimentalgruppe, d.h. mit *stereotype threat*-Manipulation resultierte unter dem Einfluss einer Annäherungs-Leistungsmotivation das vergleichsweise größte Interesse, wohingegen bei Vermeidungs-Leistungszielen oder ohne Induktion einer Zielorientierung ein stark vermindertes Interesse von Frauen festgestellt wurde. Diese

Ergebnisse deuten an, dass die Probandinnen unter Stereotypenbedrohung und ohne spezifische Induktion einer Leistungszielorientierung spontan Vermeidungs-Leistungsziele ausbilden, die wiederum mit einem verringerten Interesse einhergehen (Smith et al., 2007, *Study 1*). Eine weitere Studie ergänzt diese Ergebnisse: Werden Annäherungs- vs. Vermeidungs-Leistungsziele vor der Bearbeitung des Computer-Fähigkeitstests vorgegeben, zeigte sich unter Berücksichtigung der generellen Leistungsmotivation der Probandinnen, dass hochgradig leistungsmotivierte Versuchsteilnehmerinnen ein größeres Interesse und eine stärkere Vertiefung in die Aufgabenbearbeitung aufwiesen, wenn Annäherungs-Leistungsziele induziert wurden. Die Induktion von Vermeidungs-Leistungszielen wirkt sich hingegen ungünstig auf die beiden abhängigen Variablen aus. Probandinnen mit einer sehr geringen, generellen Leistungsmotivation zeigen konträr nach der Vorgabe eines Vermeidungs-Leistungsziels ein verstärktes Interesse und eine stärkere Involviertheit in die Aufgabenbearbeitung. Diese Ergebnisse verdeutlichen das Zusammenspiel von *stereotype threat* und motivationaler Zielorientierungen sowie deren Auswirkungen auf das domänenspezifische Interesse.

Ergänzend zeigen Steele & Ambady (2006, *Studien 1a & 1b*), dass die Einstellungen und das Interesse von Frauen in einer mathematisch-technischen Domäne nach einer subtilen Aktivierung von Geschlechterstereotypen im Vergleich zu einer Kontrollgruppe negativer ausfallen bzw. vermindert werden. Davies, Spencer, Quinn & Gerhardstein (2002) zeigen diesbezüglich, dass Frauen, denen Fernsehsequenzen mit stereotyp weiblichen Tätigkeitsbereichen gezeigt werden, sich weniger für stereotypendiskonforme Tätigkeitsbereiche interessieren, wohingegen Frauen nach der Darstellung von Fernsehsequenzen mit Inhalten, die gegen gängige Tätigkeits-Stereotype bei Frauen sprechen, ein vergleichbares Interesse an mathematisch-naturwissenschaftlichen und „verbalen“ Tätigkeiten (Journalismus, Sprach- und Kommunikationswissenschaft) berichteten. Diese Ergebnisse liefern weitere Evidenz für die Annahme, dass durch *stereotype threat* nicht nur Leistung, sondern auch Interesse als motivationale Variable beeinflusst wird. Zusätzlich wird deutlich, dass die resultierenden Auswirkungen nicht nur kurzfristige Effekte darstellen, sondern auch langfristige Konsequenzen haben können (z.B. Berufswahl).

Fogliati & Bussey (2013) weisen durch ihre Befunde darauf hin, dass eine Stereotypenbedrohung auch dazu führen kann, dass die Motivation zur *Veränderung* wahrgenommener Leistungsnachteile beeinträchtigt wird. In diesem Experiment wurden Frauen und Männer gebeten, einen standardisierten Mathematiktest durchzuführen, bei welchem sich typischerweise Geschlechterunterschiede zeigen (*threat*-Bedingung) vs. keine Geschlechterunterschiede festzustellen waren (*no threat*-Bedingung). Im Anschluss an die Bearbeitung des Tests erhielten die Teilnehmer/-innen zufällig positives oder negatives Feedback (d.h. das Feedback orientierte sich nicht an der tatsächlichen Leistung, sondern wurde randomisiert zugewiesen). Abschie-

ßend wurde die Möglichkeit gegeben, an mathematischen Fortbildungskursen teilzunehmen und die Teilnahmebereitschaft hierzu wurde abgefragt. Es zeigte sich, dass Probandinnen unter *stereotype threat* schlechtere Leistungen erzielten, als Probandinnen der Kontrollgruppe oder Männer in beiden Bedingungen. Zusätzlich wiesen Probandinnen unter Stereotypenbedrohung nach negativem Feedback eine deutlich verringerte Motivation auf, an den o.g. Fortbildungsangeboten teilzunehmen, um somit den angeblich festgestellten Kompetenzunterschieden aktiv entgegenzuwirken – für Männer kann ein derartiger Effekt nicht nachgewiesen werden. Diese Ergebnisse deuten darauf hin, dass die Problematik von *stereotype threat* möglicherweise durch die Fähigkeit einer Stereotypenbedrohung, die Motivation zur Verbesserung entsprechender Leistungsnachteile zu unterdrücken, verstärkt werden kann.

Zur Beeinflussung von Attributionen unter *stereotype threat* existieren vergleichsweise wenige empirische Studien, welche bisher auch keine eindeutige Aussage über die Auswirkungen zulassen. Koch et al. (2008) führten hierzu eine experimentelle Studie durch, in welcher die Probandinnen und Probanden eine computerbezogene Aufgabe bearbeiten sollten. Diese Aufgabe wurde je nach Gruppe so dargestellt, dass a) Geschlechterunterschiede zugunsten von Männern vorliegen (*stereotype threat*), b) Frauen typischerweise eine bessere Leistung bei der Aufgabe zeigen (positive Bedingung) oder c) keinerlei geschlechtsbezogene Informationen gegeben wurden (Kontrollbedingung). In allen drei Bedingungen war die Lösung der Aufgabe unmöglich. Die Ergebnisse zeigen, dass Probandinnen unter Stereotypenbedrohung vorzugsweise internale Attributionen formulierten, wohingegen männliche Probanden hauptsächlich externale Ursachen zur Erklärung heranzogen. In den beiden verbleibenden Bedingungen konnten keine signifikanten Unterschiede in den Attributionen festgestellt werden. Diese Ergebnisse sprechen dafür, dass unter *stereotype threat* in der Hauptsache selbstwertdienliche und nicht motivationsfördernde Attributionsmuster vorherrschen bzw. durch die Stereotypenbedrohung begünstigt werden (siehe auch: Kiefer & Shih, 2006). Diese Ergebnisse werden durch Frost (2011) und Kiefer (2006, *Study 3*) konterkariert: In beiden Studien wurde der Einfluss von *stereotype threat* auf Kausalattributionen untersucht, beide Studien finden keine signifikante Beeinflussung der Attributionsmuster durch eine Stereotypenbedrohung. Zusammenfassend wird ersichtlich, dass die wenigen, existierenden Befunde keine eindeutigen Schlussfolgerungen zulassen. Die Ergebnisse von Koch et al. (2008) sprechen für eine Beeinflussung von Kausalattributionen durch *stereotype threat*. Hinzu kommt, dass zusätzliche Befunde auf die Wichtigkeit von Attributionen in scheinbar nachteiligen Leistungssituationen hinweisen. So zeigen Wilson & Linville (1985) die leistungsförderliche Wirkung von Attributionsmustern, die bei Misserfolg eine internal-stabile Ursachenzuschreibung vermeiden. Im Kontext von *stereotype threat* kann daraus abgeleitet werden, dass negative Effekte und Erwartungen unter einer Stereotypenbedrohung kontinuierlich abgemildert werden können, wenn tatsächlicher Misserfolg nicht auf internal-stabile Ursachen

zurückgeführt wird, sondern vielmehr Erfolg derart attribuiert wird (vgl. ferner McIntyre, Paulson, Taylor, Morin & Lord, 2011). Die Befunde von Frost (2011) und Kiefer (2006) widersprechen diesen Ergebnissen und Ableitungen allerdings zumindest teilweise. Demzufolge scheint es zum jetzigen Zeitpunkt auch im Hinblick auf die quantitative Anzahl verfügbarer Studien (noch) nicht möglich, eine umfassende Bewertung vorzunehmen.

Ein Einfluss von *stereotype threat* auf volitionale Prozesse ist zum Zeitpunkt der Entstehung dieser Arbeit bisher nicht Untersuchungsgegenstand publizierter, empirischer Arbeiten gewesen, weshalb hier ebenfalls zum jetzigen Zeitpunkt keine schlüssigen und abgesicherten Aussagen über mögliche Zusammenhänge zwischen diesen Konstrukten abgeleitet werden können.

Abschließend weisen Forbes & Schmader (2010) mit ihren Forschungsbefunden darauf hin, dass es sich bei der Beeinträchtigung der Motivation unter *stereotype threat* keineswegs um einen unumkehrbaren Effekt handelt, sondern die Motivation durch ein Re-Training wieder positiv beeinflusst werden kann. In Studie 1 wurde ein auf dem impliziten Assoziationstest (IAT, Greenwald, McGhee & Schwartz, 1998) basiertes Trainings absolviert, in welchem die Probandinnen lernten, sich selbst mit einer positiven Einstellung gegenüber der Mathematik zu assoziieren („Ich mag“ & „Mathematik“). In der Kontrollbedingung wurde eine gegenteilige Beziehung („Ich mag nicht“ & „Mathematik“) gelernt. Anschließend wurde in allen Studien eine Auswahl-aufgabe präsentiert, in welcher entweder ein mathematisches oder ein nicht-mathematisches Problem gewählt werden konnte. Zur Erfassung der Arbeitsgedächtnisspanne wurde ebenfalls in allen Teilstudien der *reading span test* (Daneman & Carpenter, 1980) eingesetzt, die Erfassung der mathematischen Leistungsfähigkeit wurde über einen Grundlagentest mathematischer Fähigkeiten vorgenommen. Es zeigte sich, dass Probandinnen in der Experimentalbedingung signifikant mehr mathematische Probleme zur Bearbeitung auswählten und diese deutlich länger betrachteten, als Probandinnen der Kontrollgruppe. Auf Leistung, Arbeitsgedächtnisspanne und Einstellung gegenüber der stereotypisierten Domäne Mathematik konnte (noch) kein Einfluss der Trainingsbedingungen nachgewiesen werden. Neben diesem ersten Hinweis auf eine erhöhte Motivation in der Experimentalbedingung blieb unklar, inwieweit sich dieses Befundmuster unter Stereotypenbedrohung darstellt. Hierzu wurden in Studie 2 auch männliche Probanden mit einbezogen und zufällig einer von drei Bedingungen zugewiesen: IAT mit dem Assoziationspaar „Ich mag“ und „Mathe“ mit *stereotype threat* vs. ohne *stereotype threat* oder „Ich mag nicht“ und „Mathe“ ohne Stereotypenbedrohung. Es folgte eine Abfrage der Motivation in geschlechtergemischten Gruppen. Darauf aufbauend wurde in den entsprechenden Untersuchungsgruppen die Stereotypenbedrohung darüber aktiviert, dass Probandinnen unter Anwesenheit eines männlichen Versuchsteilnehmers zunächst ihr Geschlecht angeben mussten. Ferner wurde der folgende Test als „diagnostischer Mathematiktest“ vs. als Problemlöseaufgabe eingeführt. Es

zeigte sich, dass in der Gruppe der Frauen mit dem „Ich mag“ & „Mathe“-IAT und einer Stereotypenbedrohung die Zeit der Aufgabenbearbeitung höher ausfiel, als in allen Vergleichsgruppen mit Ausnahme der Frauen im stereotypenkonformen Training ohne *stereotype threat*. Für Männer konnten keinerlei signifikante Effekte festgestellt werden. Es konnte also gezeigt werden, dass ein Retraining zur Assoziation besserer Einstellungen gegenüber der stereotypisierten Domäne in einer erhöhten Motivation die Aufgabenbearbeitung betreffend resultiert und dieser Effekt ganz besonders unter Stereotypenbedrohung zum Tragen kommt. In Studien 3 und 4 wurde unter Hinzunahme einer weiteren Trainingsmethode, die das Geschlechtsstereotyp *per se* anspricht, ergänzend zu den motivationalen Variablen gezeigt, dass ein IAT, welcher „Frauen = gut in Mathe“ assoziiert die Arbeitsgedächtniskapazität insbesondere unter *stereotype threat* erhöhen und als Mediator für eine verbesserte Mathematik-Testleistung herangezogen werden kann. Dies macht deutlich, dass es sich bei den die Motivation (und Leistung) hemmenden Einflüssen von *stereotype threat* keineswegs um unveränderliche Zusammenhänge handelt, sondern dass diese schon mithilfe einer vergleichsweise unaufwändigen Intervention zum Besseren beeinflusst werden können.

Die meisten Erkenntnisse zu Einflüssen von Stereotypenbedrohungen auf die Motivation liefern Studien zur Leistungsmotivation (Rheinberg, 2006), genauer zu Zielorientierungen: Dweck & Leggett (1988) untersuchten mit Versuchspersonen im Kindesalter den Zusammenhang von Zielorientierungen und daraus resultierenden Verhaltensmustern und wiesen darauf hin, dass *Lernziele* unabhängig von der derzeit wahrgenommenen Fähigkeit in ein Verhalten münden, welches durch *mastery*, also Kontrolle und Bewältigung, gekennzeichnet ist. Dass auch eine niedrige Wahrnehmung der eigenen Kompetenz in Verbindung mit Lernzielen zu *mastery* führt, wird damit begründet, dass bei dieser Zielorientierung der Kompetenzzuwachs im Vordergrund steht und daher das aktuelle Kompetenzniveau nur eine untergeordnete Rolle spielt. Für Personen mit *Leistungszielorientierung* sagen die Autorinnen *mastery* ebenfalls voraus, wenn das durch die Person wahrgenommene, aktuelle Fähigkeitsniveau hoch ausfällt. Im Falle der Wahrnehmung eines geringen Kompetenzniveaus wird ein negatives Verhaltensmuster postuliert, welches durch Hilflosigkeit und Vermeidungsverhalten geprägt ist (Dweck & Leggett, 1988). In einer experimentellen Überprüfung wurden Lern- und Leistungsziele induziert und die wahrgenommene Kompetenz (durch Feedback über eine Testaufgabe) der Probanden/-innen manipuliert, es resultierte somit ein 2 x 2-Design (Lern- vs. Leistungszielorientierung / aktuelle Kompetenz hoch vs. gering). Der Erfolg der Induktion eines Lern- bzw. Leistungsziels wurde überprüft, indem die Probanden auswählen sollten, ob sie lieber eine Aufgabe bearbeiten wollen, die einem Lernziel gleich kommt (hohe Aufgabenschwierigkeit verbunden mit dem Risiko, die Aufgabe nicht lösen zu können, aber mit substanziellem Lernpotential) oder die einer Leis-

tungszielorientierung entspricht (Aufgabe, die vergleichsweise einfach zu lösen ist, dafür aber kein nennenswertes Lernpotential bietet). Durch die Methode des lauten Denkens während der Problemlösung sollten alle Probanden/-innen ihre Gedanken und Lösungsschritte verbalisieren, welche aufgezeichnet und kategorisiert wurden. Ferner wurden die Versuchspersonen gebeten, verschiedene Aufgabenstellungen auszuwählen – welche nach dem Grad von Schwierigkeit und Herausforderung variierten – um den folgenden Testabschnitt entsprechend den gemachten Vorgaben zusammenstellen zu können. (*concept formation task*; Diener & Dweck, 1978).

Aus der Analyse der so gewonnenen Daten wurde deutlich, dass bei Induktion eines *Lernziels* unabhängig vom aktuellen Fähigkeitsniveau ein herausforderndes Aufgabenniveau gewählt wurde und mehr verbale Äußerungen vorkamen, die auf *mastery* hinweisen. Nach Induktion eines *Leistungsziels* war die Auswahl der Aufgabe und der verbalen Indikatoren für ein Verhaltensmuster in hohem Maße von der wahrgenommenen Kompetenz abhängig. Wurde durch positives Feedback eine hohe Kompetenz antizipiert, wurden vermehrt schwierige, herausfordernde Aufgaben gewählt. Bei negativem Feedback und damit einhergehend einer Wahrnehmung der aktuellen Fähigkeiten als gering zeigte sich die vorhergesagte Hilflosigkeit in Verbindung mit einer Vermeidung von Misserfolg, wie z.B. eine Auswahl von Aufgaben mit geringem Schwierigkeitsgrad. Dieses Befundmuster bleibt ebenso bestehen, wenn Zielorientierungen nicht induziert, sondern bestehende (reelle) Zielorientierungen untersucht werden (Bandura & Dweck, 1985).

Die bisher berichteten, empirischen Ergebnisse stellen einen wichtigen Entstehungszusammenhang für die Einbettung von *stereotype threat*-Effekten dar. Neben Befunden aus einem breiten Spektrum der Motivation zeichnet sich jedoch ab, dass der Leistungszielorientierung eine besondere Rolle bei der Erklärung negativer Auswirkungen einer Stereotypenbedrohung auf die Motivation zukommt. Zunächst basiert das Paradigma *stereotype threat* kongruent zu o.g. Befunden zur Leistungszielorientierung auf einer Wahrnehmung geringer Kompetenz. Daraus resultieren u.a. negative Affekte sowie die Motivation, mögliche Inkompetenz zu verbergen (im Sinne von *stereotype threat*: das negative Stereotyp nicht zu bestätigen). Nach Dweck & Leggett (1988) ist es unter Stereotypenbedrohung also die Einnahme einer Leistungszielorientierung abgeleitet aus den theoretischen Annahmen und gestützt durch die empirischen Ergebnisse möglicherweise eine adäquate Strategie, um das negative Stereotyp nicht zu bestätigen. Auch ist es von grundlegender Bedeutung, die Rolle von Leistungszielen in *stereotype threat*-Situationen angemessen zu berücksichtigen, da sich Individuen in Lern- und Leistungssituationen in Abhängigkeit von der eingenommenen oder zugewiesenen Zielorientierung mit diversen Aufgaben verschiedenartig auseinandersetzen, diese unterschiedlich erleben und ausführen. In diesem Zusammenhang sah Smith (2004) in der Zielorientierungsforschung ein bisher fehlendes Stück zur Erklärung der Wirkungsweise von Stereotypenbedrohung und bezog daher o.g. Befunde zur

Zielorientierung in die aktuelle *stereotype threat*-Forschung ein, was in einer Integration von Zielorientierungen im Modell des *stereotyped task engagement process* (STEP; Smith, 2004) resultierte.

3.3. Das Stereotyped Task Engagement Process-Modell

Das Stereotyped Task Engagement Process (STEP)-Modell (siehe Abbildung 2) wurde erstellt, um Auswirkungen von *stereotype threat* auf verschiedene Bereiche (nicht nur auf Leistungsebene) zu beschreiben. Der Leistungszielorientierung kommt dabei eine besondere Bedeutung zu, da bisherige Forschung zu Zielorientierungen nahelegt, dass die Interaktion mit Aufgaben, deren Wahrnehmung und die resultierende Leistung aufgrund verschiedener, zugrundeliegender Leistungsziele divergiert (Barron & Harackiewicz, 2001; Dweck & Leggett, 1988). STEP ist somit ein Modell, um Forschungsergebnisse über sowohl *stereotype threat* als auch Leistungszielorientierung zu verbinden und besser zu verstehen.

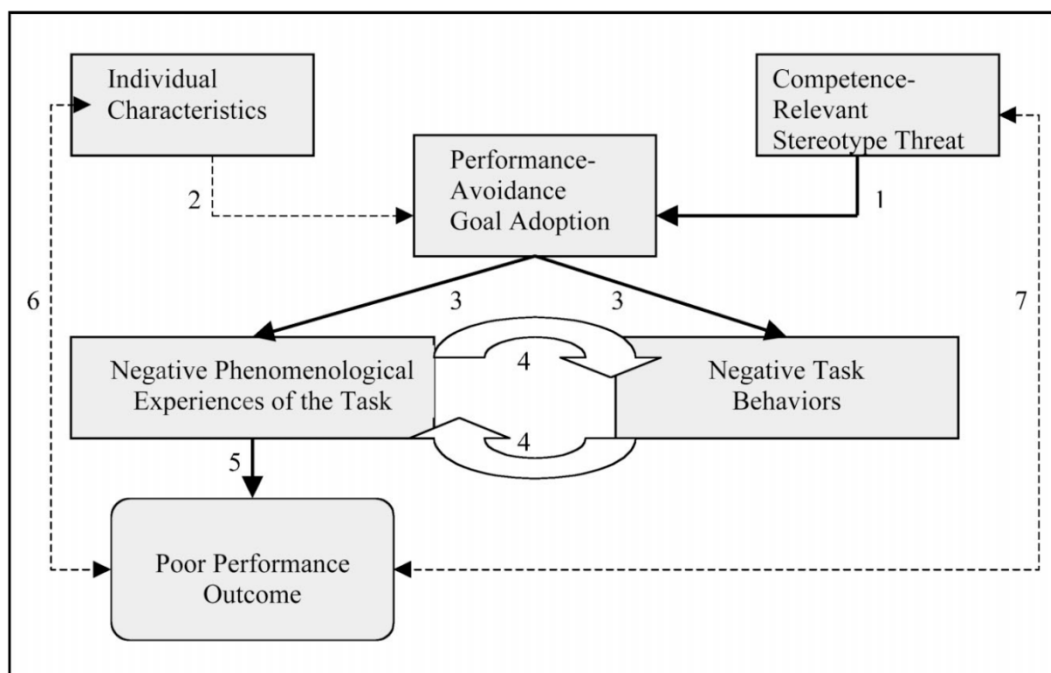


Abbildung 2: Stereotype Task Engagement Process-Modell (STEP). Aus: Smith (2004), S. 194.

Im Folgenden sollen die im Modell angenommenen Beziehungen zwischen den einzelnen Komponenten (dargestellt durch Pfeile) erläutert werden: *Pfeil 1* symbolisiert den angenommenen Zusammenhang, dass negative Informationen über Stereotype während der Ausführung einer stereotyp-relevanten Aktivität zur Annahme von *performance avoidance (PAV)-goals* führen können. Forschungsergebnisse über die Leistungszielorientierung zeigen, dass es notwendig ist, die allgemeine Reaktion einer Person auf die Leistungszielorientierung mit zu beachten, um die Effekte einer Leistungszielorientierung auf die Leistung und Motivation zu verstehen (Hara-

ckiewicz & Elliot, 1993). *Pfeil 2* beschreibt den moderierenden Einfluss von individuellen Unterschieden (z.B. Leistungsmotivation, Identifikation mit einem bestimmten Bereich) auf die Bereitschaft ein *PAV*-Ziel einzunehmen. *Pfeil 3* stellt den Einfluss von *PAV*-Zielorientierung auf phänomenologische Erlebensprozesse (Interesse, Angst, Zuversicht) und Verhalten (Anstrengung, Lernstrategien etc.), wobei das STEP-Modell davon ausgeht, dass eine Person abhängig von der Leistungszielorientierung nur bestimmte Selbstregulationsstrategien anwendet. *Pfeil 4* stellt dar, dass die phänomenologischen Erfahrungen die Interaktion mit der Aufgabe beeinflussen und diese wiederum auf die phänomenologischen Erfahrungen rückwirken. Nach den Annahmen des Modells führt somit die negative phänomenologische Erfahrung der Interaktion mit der Aufgabe zu einer Verminderung von leistungsbezogenem Verhalten (z.B. Anstrengung). Dies wirkt sich dann wiederum auf die Wahrnehmung der Aufgabe aus usw. *Pfeil 5* stellt den Einfluss der negativen, phänomenologischen Erfahrung auf die Testleistung dar. *Pfeil 6* weist auf den moderierenden Einfluss verschiedener individueller Eigenschaften (z.B. Identifikation mit der Domäne; Interesse) auf die Leistung hin und postuliert gleichzeitig auch einen Einfluss der Leistung auf die Ausgestaltung dieser individuellen Charakteristika (z.B. ist denkbar, dass fortlaufend gute Leistungen in einem Bereich die *domain identification* erhöhen). *Pfeil 7* deutet abschließend darauf hin, dass die individuelle Leistung bei einer stereotyp-relevanten Aufgabe die (In-)Validität des Stereotyps unterstützen oder untergraben kann und somit zu einer Verfestigung oder Auflockerung des leistungsrelevanten Stereotyps führen kann.

Insgesamt kann also festgehalten werden, dass das STEP-Modell von Smith (2004) davon ausgeht, dass eine kompetenzrelevante Bedrohung zur Einnahme eines Vermeidungs-Leistungsziels führt. Dies wiederum resultiert in einem negativ konnotierten Erleben der Aufgabe und in negativem, aufgabenbezogenem Verhalten, wobei beide Aspekte sich wechselseitig beeinflussen und verstärken können. Schlussendlich münden diese Prozesse in einer verminderten Leistung, medierend wirken individuelle Charakteristika der Person und die aus der Leistung resultierende Verfestigung oder Auflockerung der stereotypen Leistungsannahme.

Evidenz für das von Smith (2004) postulierte Modell kommt u.a. aus Forschungsarbeiten von Brown & Josephs (1999, *Study 1*). Es wurde ein Mathematiktest durchgeführt, welcher einmal als Test dargestellt wurde, welcher konzipiert wurde um besonders leistungsfähige Individuen zu identifizieren. In einer anderen Bedingung wurde dargelegt, der Test wäre zur Identifikation besonders leistungsschwacher Personen angelegt. Frauen, die dachten, dass sie einen Mathematiktest durchführen, der besonders die Schwachen identifiziert, erzielten schlechtere Ergebnisse als Frauen, die dachten, der Mathematiktest würde besonders starke identifizieren. Bei Männern zeigte sich ein genau umgekehrtes Ergebnismuster: Hier war die Leistung bei Männern in der Bedingung „leistungsstarke Personen identifizieren“ gegenüber Männern in der Bedingung „leistungsschwache Personen identifizieren“ verringert. In dieser Untersuchung wurde

stereotype threat zwar nicht explizit manipuliert oder gemessen, die Autoren interpretieren ihre Ergebnisse dennoch im Sinne des *stereotype threat* mit dem Argument, das Geschlechterstereotyp über Mathematik wird von Frauen und Männern wahrgenommen (im Sinne einer subtilen Stereotypenaktivierung durch die reine Angabe des Geschlechts). Zusätzlich ist die Manipulation der Besorgnis vergleichbar mit der Manipulation von *performance approach- (PAP)-goals* und *PAV-goals* von Elliot, McGregor & Gable (1999). Obwohl Brown & Josephs (1999) keinen expliziten Bezug auf Leistungsziele nehmen oder eine klare Verbindung hierzu herstellen, so deuten die Ergebnisse dennoch darauf hin, dass Individuen in Konfrontation mit Stereotypen spontan eine Leistungszielorientierung einnehmen, welche wiederum die Leistungsergebnisse beeinflusst (Smith, 2004).

Elliot & Harackiewicz (1996) liefern weitere Evidenz für das Modell. Wenn Versuchspersonen vor Durchführung der Aufgabe (in diesem Fall eine Puzzle-Aufgabe) ein *PAV-goal* (zeigen zu können, dass man kein schlechter Puzzler ist) vorgegeben wird, dann resultiert eine geringere Motivation im Vergleich zu Versuchspersonen, denen ein *PAP-goal* (die Möglichkeit, zeigen zu können, dass man ein guter Puzzler ist) induziert wurde. Somit zeigen auch diese Ergebnisse den vorhergesagten Zusammenhang zwischen der Induktion eines *PAV-Ziels* und der resultierenden (schwächeren) intrinsischen Motivation. Der Zusammenhang mit *stereotype threat* bleibt zunächst nur theoretisch herzustellen: Unter Stereotypenbedrohung dominiert das Ziel, die Darstellung zu vermeiden in einer (durch Stereotype bedrohten) Disziplin schlecht abzuschneiden, d.h. das Stereotyp möglicherweise zu bestätigen (= *PAV-Ziel*). Hieraus resultiert eine Untergrabung der intrinsischen Motivation und resultiert schlussendlich in schlechteren Leistungsergebnissen. Somit kann auch diese Studie erste Evidenz für die Gültigkeit des Modells geben.

Eine weitere, direktere Prüfung fand durch Smith (2006) statt. In zwei empirischen Studien wurden die Annahmen geprüft, dass Frauen unter *stereotype threat* verstärkt eine Vermeidungs-Leistungszielorientierung annehmen sowie dass dieses Befundmuster auch im Vergleich mit Männern auftritt und ergänzend die Vermeidungs-Leistungsziele zu einer geringeren Leistungserwartung führen. In Studie 1 wurde nach der Erfragung demografischer Daten die Identifikation mit der Leistungsdomäne gemessen und ein Zeitungsartikel vorgelegt, welcher explizit auf Geschlechterstereotype zuungunsten von Frauen in Mathematik hinwies. Darauf aufbauend wurde in der Experimentalbedingung erwähnt, dass die im Zeitungsartikel angesprochenen Unterschiede auch in dieser Untersuchung festgestellt wurden (konforme Bedingung), in der Kontrollgruppe hingegen wurde darauf hingewiesen, dass in der folgenden Untersuchung keinerlei derartige Geschlechtsunterschiede gefunden werden konnten (*nullified*-Bedingung). Darauf folgend wurden die Leistungsziele (*performance approach* und *performance avoidance*) gemessen. Ein Mathematiktest, wie eingangs angekündigt, war nicht Bestandteil dieser Studie. Es zeigte sich, dass Frauen in der Experimentalbedingung sich signifikant häufiger einer Vermeidungs-

Leistungszielorientierung zuwenden, als die Probandinnen der Kontrollbedingung. Hinsichtlich der Annäherungs-Leistungsziele wurden keine signifikanten Unterschiede zwischen den Gruppen festgestellt. Da nur Frauen an der Studie teilnahmen bleibt unklar, wie sich Frauen und Männer in ihren Leistungszielen unterscheiden. Um diese Frage beantworten zu können wurde Studie 2 konzipiert. An dieser Studie nahmen Männer und Frauen teil, der Ablauf entsprach dem aus Studie 1. Abweichend zu Studie 1 wurde allerdings nur eine stereotypenkonforme Bedingung realisiert, zusätzlich wurde die Erfolgserwartung über eine Ein-Item-Messung mit erhoben. Regressionsanalytisch wurde nachgewiesen, dass nur Vermeidungs-Leistungsziele bedeutsam vom Geschlecht der Versuchsperson vorausgesagt werden, was die Annahme bestätigt, dass in dieser Studie die bedrohte Gruppe (Frauen) eher eine Vermeidungs-Leistungszielorientierung annimmt, als die Fremdgruppe (Männer). Diese Ergebnisse schließen zwar nicht die Möglichkeit aus, dass Frauen allgemein eher Vermeidungs-Leistungsziele annehmen als Männer. Die Theorie zu Leistungszielorientierungen unterstützt diese Erklärung allerdings nicht (z.B. Pintrich, 2000). Zusätzlich wurde gezeigt, dass mit einer steigenden Annahme einer Vermeidungs-Leistungszielorientierung eine sinkende Leistungs-/ Erfolgserwartung einhergeht. Erneut konnte hinsichtlich der Annäherungs-Leistungsziele kein bedeutsamer Befund festgestellt werden. Die Autorin interpretiert dies so, dass trotz des gleichzeitigen Vorhandenseins von Annäherungs-Leistungszielen (im Sinne eines *multiple goal*-Ansatzes; Barron & Harackiewicz, 2001) der Einfluss der Vermeidungs-Leistungsziele ausreicht, um mögliche positive Einflüsse durch die Annäherungs-Leistungsziele zunichte zu machen.

Insgesamt wäre es für eine umfangreiche Prüfung des Modells hilfreich gewesen, auch die Auswirkungen auf der Leistungsebene mit zu betrachten. Nichtsdestotrotz wurde oft nachgewiesen, dass geringere Erwartungen zu einer Verschlechterung der Leistung (Bandura, 1997) und einem niedrigeren Ausmaß an intrinsischer Motivation (Ferrer-Caja & Weiss, 2002) führen können. Im Hinblick darauf unterstützen die soeben erläuterten Befunde das Modell von Smith (2004).

Des Weiteren prüften Smith et al. (2007) in drei umfangreichen Studien die Gültigkeit des Modells. Die zwei bereits auf Seite 60 ausführlich erläuterten Untersuchungen zeigten, dass Frauen mit einer hohen Leistungsmotivation unter Stereotypenbedrohung ein verringertes Interesse gegenüber einer technisch-mathematischen Aufgabe zeigten, wenn ihnen ein Vermeidungs-Leistungsziel induziert wurde oder wenn keinerlei Zielorientierung vorgegeben wurde, was wiederum dafür spricht dass hier spontan Vermeidungs-Leistungsziele aktiviert wurden (*Study 1*). Weiterhin wurde gezeigt, dass hochgradig leistungsmotivierte Probandinnen ein verstärktes Interesse und eine stärkere Vertiefung in die Aufgabe aufweisen, wenn zuvor ein *performance approach-goal* induziert wurde resp. Interesse und Grad der Vertiefung in eine Aufga-

be sinken, wenn ein *performance avoidance-goal* zugewiesen wird (*Study 3*). Zusätzlich wurden in Studie 2 Probandinnen keine Zielorientierungen zugewiesen, sondern das spontane Auftreten von Zielorientierungen wurde untersucht. Die Versuchspersonen wurden erneut gebeten, eine schwierige Mathematikaufgabe zu lösen und dabei ihre Gedanken während der Aufgabebearbeitung zu verbalisieren, zusätzlich wurde neben einer Gruppe mit Stereotypenbedrohung und einer weiteren ohne *stereotype threat* noch eine dritte Gruppe hinzugefügt, in welcher keinerlei Angaben zu Stereotypen gemacht wurden. Es zeigte sich, dass hoch leistungsmotivierte Personen unter *stereotype threat* spontan mehr Gedanken äußerten, die einer Vermeidungs-Leistungszielorientierung gleichkommen, als Personen mit geringerer Leistungsmotivation oder der Kontrollbedingung. In der Versuchsgruppe, in welcher *keinerlei* Angaben über Stereotype gemacht wurden, zeigten hoch leistungsmotivierte Probandinnen exakt gleiche Effekte (mehr Äußerungen, die auf eine Vermeidungs-Leistungszielorientierung hinweisen), wie bei hoch leistungsmotivierten Versuchspersonen in der Bedrohungsbedingung. Dies deutet darauf hin, dass leistungsbezogene Stereotype in dieser Studie auch ohne explizite Erwähnung des Faktors Geschlecht aktiviert wurden.

Evidenz für das STEP-Modell (Smith, 2004) ergibt sich somit insgesamt in der Hauptsache aus dem Nachweis, dass *stereotype threat* zu einer Vermeidungs-Leistungszielorientierung führt sowie das eine (zugewiesene und spontan aufgetretene) Vermeidungs-Leistungszielorientierung zu einer negativen Wahrnehmung der Aufgabe (Interesse, wahrgenommene Kompetenz) und negativem, aufgabenbezogenem Verhalten (*task engagement / task absorption*) führt, was nach den Annahmen des Modells zu erwarten wäre. Nichtsdestotrotz treten differenziertere Befunde auf, u.a. scheinen die gerade beschriebenen Zusammenhänge ausschließlich für Personen zu gelten, die eine hohe Leistungsmotivation aufweisen, denn bei Personen mit einer geringen Leistungsmotivation zeigt sich ein umgekehrtes Befundmuster. Diese moderierenden Einflüsse der Leistungsmotivation werden so im Modell nur angedeutet und den „individuellen Charakteristika“ zugeordnet, es scheint jedoch ein expliziterer Einbezug in das Modell notwendig zu sein. So kann durch eine verstärkte Berücksichtigung der Leistungsmotivation erklärt werden, warum unter Stereotypenbedrohung Individuen mit einer geringen Leistungsmotivation spontan ein Annäherungs-Vermeidungsziel einnehmen. Bei Personen mit einer niedrigen Leistungsmotivation scheint der Aspekt Leistung kein zentraler Bestandteil des domänenspezifischen Selbstkonzepts für eine vorliegende Aufgabe darzustellen, weshalb potentieller und durch *stereotype threat* getriggelter, antizipierter Misserfolg weniger selbstwertbedrohlich wirkt und so zur Wahl vergleichsweise „risikoreicherer“ *performance approach*-Zielorientierungen führen kann (Smith et al., 2007).

Zusätzlich schlagen die Autoren/-innen vor, das STEP-Modell (Smith, 2004) um einen Faktor „Interesse“ zu erweitern, um mögliche Zusammenhänge von Stereotypenbedrohung, Interes-

se und Leistung besser zu berücksichtigen. Es wird angenommen, dass die Zusammenhänge zwischen Interesse und Leistung zeitlich variabel und abhängig von der Leistungszielorientierung sind, was im Modell so noch nicht angemessen berücksichtigt wird. Drittens wird als Konsequenz aus den Ergebnissen für das Modell abgeleitet, dass die *ausschließliche* Fokussierung auf Leistungszielorientierungen zugunsten eines verstärkten Einbezugs anderer Typen von Zielen (z.B. interpersonale Ziele) aufgegeben werden sollte.

Abschließend wird gefordert, dieses nunmehr erweiterte Modell verstärkt auf einen Anwendungskontext zu applizieren, in dem von Stereotypen bedrohte, aber trotzdem erfolgreiche Personengruppen zum Untersuchungsgegenstand werden (z.B. erfolgreiche, weibliche Studierende in MINT-Fächern), um feststellen zu können, welche motivationalen Mechanismen trotz *stereotype threat* zum Erfolg führen. Insgesamt liefern die Befunde von Smith et al. (2007) zu einem großen Anteil Evidenz für das postulierte Modell von Smith (2004), wenngleich an einigen Stellen noch Anpassungen vorgeschlagen werden, um den prädiktiven Wert des Modells sukzessive zu erhöhen.

Thoman et al. (2013) schlagen ein überarbeitetes Modell, das *Motivational Experience Model of Stereotype Threat* vor. Dieses Modell, was die Grundannahmen von Smith (2004) beinhaltet und an neuere Forschungsergebnisse angepasst hat, geht davon aus, dass eine Vielzahl motivationaler Variablen, welche eine zentrale Bedeutung für die Akquise akademischer Kompetenzen aufweisen, durch Stereotypenbedrohung beeinflusst werden. Insbesondere ein Einbezug einer Ebene der Selbstregulation wird als wichtige Ergänzung in diesem Modell eingeführt. Des Weiteren wird zusätzlich ein soziales Zugehörigkeitsmotiv („*sense of belonging*“) und die intrinsische Motivation, resp. Interesse neben Leistungszielen in das überarbeitete Modell aufgenommen. Die Verknüpfungen aller durch *stereotype threat* beeinflussten Komponenten des Modells (also Interesse und Zugehörigkeitsmotiv mit Ausnahme der motivationalen Selbstregulation, s.u.) wurden aus vorherigen Forschungsergebnissen abgeleitet und weisen daher eine substantielle empirische Evidenz auf. Dem Modell folgend beeinflusst *stereotype threat*, wie sich eine Person an eine Aufgabe annähert und die Aufgabe wahrnimmt (analog zu den Annahmen des Modells von Smith, 2004). Das phänomenologische Erleben wird hingegen in diesem Modell differenzierter betrachtet und besteht aus zwei Komponenten: Interesse, resp. der intrinsischen Motivation, und dem sozialen Zugehörigkeitsmotiv. Nach dem Modell schlägt sich eine Stereotypenbedrohung mit hoher Wahrscheinlichkeit in einem verminderten Interesse und/oder einer geringeren, wahrgenommenen sozialen Eingebundenheit nieder. Es wird somit deutlich, dass *stereotype threat* eine oder beide Komponenten stören kann. Dies lässt selbstregulatorische Prozesse notwendig werden, welche ebenfalls durch die Stereotypenbedrohung nur eingeschränkt zur Verfügung stehen. Es wird davon ausgegangen, dass die Stereotypenbedrohung auf direktem Wege auf die Leistungszielorientierung, die phänomenologische Wahrnehmung der Aufgabe

(Interesse, *sense of belonging*) und die Selbstregulation motivationaler Prozesse einwirkt und hierdurch die negativen *Outcomes* auf motivationaler und Leistungsebene, einschließlich der Identifikation mit der Domäne maßgeblich erklärt werden können.

Die Zusammenhänge zwischen *stereotype threat*, Selbstregulation und den anderen Komponenten des Modells bleiben abschließend noch auf empirische Evidenz zu überprüfen. Diesbezüglich konnte in weiteren Untersuchungen gezeigt werden, dass wenn stigmatisierte Lerner/-innen wahrnehmen, dass eine Regulation im Sinne von einer Erhöhung der Motivation, resp. des Interesses, im Angesicht von *stereotype threat* zu keiner Leistungsverbesserung führt (und damit die Bedrohung immanent wird, das Stereotyp zu bestätigen), kann dies eine Zerstörung der wahrgenommenen Kosten-Nutzen-Relation für die Regulation des Interesses zur Folge haben und somit dazu beitragen, die Regulation der intrinsischen Motivation/des Interesses unter Stereotypenbedrohung aktuell und auch künftig einzuschränken (Hidi & Renninger, 2006).

Stereotype threat führt ferner zu einer Wahrnehmung sozialer Ausgrenzung (Richman & Leary, 2009). Die Verletzung von stereotypen, prototypischen Annahmen über die Leistung einer Person in einer bestimmten Domäne kann ebenfalls beeinträchtigend auf die Wahrnehmung sozialer Verbindungen und das Selbstbild wirken (Kessels, 2005). Eine regulative Strategie, mit dieser antizipierten sozialen Ausgrenzung umzugehen ist die Möglichkeit, sich den existierenden Annahmen und Leistungsnormen (z.B. „Mädchen können kein Mathe“) zu unterwerfen (Smith & Lewis, 2009) oder sich von stigmatisierten, identitätsrelevanten Aspekten zu entfernen (z.B. Frauen, welche sich von femininen Aspekten ihrer Identität distanzieren, um sich in das Mathe = männlich Stereotyp besser einzupassen; Major, Quinton, McCoy & Schmader, 2000). Weiterhin ist möglich, sich von der Domäne (z.B. Mathematik) zu entfernen, d.h. die *domain identification* herunter zu regulieren (Schmader, Major & Gramzow, 2001). Wie deutlich wird, legen die zitierten Studien nahe, dass die Möglichkeiten zur Regulation eines Gefühls verminderter sozialer Eingebundenheit, welches durch *stereotype threat* hervorgerufen wird vornehmlich negative Auswirkungen auf die phänomenologische Wahrnehmung der Aufgabe aufweisen (mit Ausnahme der Befunde von Major et al. (2000), welche auf negative, identitätsrelevante Konsequenzen hinweisen) und somit eher Distanz, als Annäherung zur Aufgabe erzeugen – und somit auch mit negativen Konsequenzen auf Leistungsebene verbunden sind.

Abschließend bleiben selbstregulatorische Auswirkungen auf die Leistungsmotivation zu berücksichtigen. Neben den in diesem Abschnitt bereits ausführlich erläuterten Effekten auf die Leistungszielorientierung, d.h. dass unter *stereotype threat* vornehmlich Leistungsziele herausgebildet werden, um sich vor möglichem Misserfolg abzuschirmen (Smith, 2004, 2006; Smith et al., 2007) zeigt Keller (2002) u.a., dass Personen unter Stereotypenbedrohung verstärkt auf *self-handicapping* als Gegenregulationsstrategie zurückgreifen, was negative Auswirkungen für den eigentlichen Prozess der Aufgabenbearbeitung nach sich zieht.

Insgesamt findet sich somit eine große, empirische Evidenz, welche für die Gültigkeit der Vorhersagen auf Basis des erweiterten Modells von Smith (2004) durch Thoman et al. (2013) spricht. Zusammenfassend lässt sich festhalten, dass die o.g., ausgewählten empirischen Studien und Modelle (Smith, 2004, Thoman et al., 2013) nahelegen, dass bei der Analyse von Effekten durch Stereotypenbedrohung unbedingt auch motivationale Prozesse einbezogen werden müssen. Ferner weist die empirische Datenlage darauf hin, dass dem Konzept der motivationalen Zielorientierung eine besondere Rolle zukommt, d.h. dass die Bedrohung durch ein leistungsrelevantes Stereotyp in der Einnahme einer Vermeidungs-Leistungszielorientierung münden kann, welche sich ungünstig auf intrinsische Motivation und Leistung auswirkt. Eine Applikation dieser Befunde auf die Problemstellung der vorliegenden Arbeit soll im folgenden Abschnitt vorgenommen werden.

3.4. Anwendung auf die Problemstellung

Lernmotivationen beeinflussen die Handlungen und die Strategien, welche zum Lernen als Wissens- und Kompetenzerwerb eingesetzt werden. Bestimmte motivationale Lagen wirken sich dabei förderlich oder hemmend auf den Lernzuwachs aus. So können durch eine Attribution von Erfolgen auf internal-stabile Ursachen oder die Verfolgung einer Lernzielorientierung langfristig erfolgreiche Lernprozesse gesichert werden (Utman, 1997; Weiner, 1986).

Im Kontext der Fragestellung der vorliegenden Arbeit bleibt nun zu klären, ob und wenn ja inwiefern sich der öffentliche Diskurs über „Jungen als Bildungsverlierer“ auf lernmotivationale Variablen auswirkt und dazu führt, dass die Auseinandersetzung von Jungen mit einem Lerngegenstand beeinflusst und damit einhergehend das Lernpotential nicht voll ausgeschöpft wird. Es bleibt also zu klären, ob die öffentliche Debatte hier gleichsam als Bedrohung wirkt und neben leistungsmindernden Effekten auch eine Beeinträchtigung der prospektiven Motivation mit sich bringt. Erste Befunde deuten an, dass durch die Wirkung von *stereotype threat* lernförderliche, motivationale Lagen (z.B. Lernzielorientierung; Dweck & Leggett, 1988) gehemmt werden und damit verbunden der Kompetenzzuwachs negativ beeinflusst wird. Insbesondere der Leistungszielorientierung kann hierbei eine besondere Bedeutung zukommen, da nach Smith (2004, 2006) durch Stereotypenbedrohung eine Ausbildung von *performance avoidance-goals* angenommen werden kann, welche wiederum zu einer unorganisierten Verwendung von Lernstrategien (Elliot et al., 1999), erhöhter Angst und geringerem Interesse einhergehen, die schlussendlich in einer schlechteren Leistung münden (Smith, 2004).

Eine Anwendung speziell auf die „Jungenproblematik“ ist derzeit noch nicht erfolgt, was im Rahmen dieser Arbeit geändert werden soll. Die o.g. Befunde bieten somit eine wichtige Perspektive bei der Analyse des Bildungserfolgs von Jungen: Es gibt möglicherweise Auswirkungen

durch den öffentlichen Diskurs über das angebliche Scheitern von Jungen im Sinne von *stereotype threat*, welche jenseits von Leistung die Lernprozesse von Jungen nachhaltig beeinflussen können und somit als Erklärungsmöglichkeit für den vergleichsweise nachlassenden Bildungserfolg von Jungen herangezogen werden können. So soll geprüft werden, ob eine *stereotype threat*-Situation zum vermehrten Einsatz von Vermeidungs-Leistungszielen mit den damit einhergehenden, lernhinderlichen Konsequenzen führt, was durch die theoretischen Annahmen und empirischen Befunde des vorangegangenen Abschnittes nahelegt. Es sollen somit die zentralen Befunde zu Stereotypenbedrohung und Motivation, welche in der Hauptsache auf Frauen im MINT-Bereich fokussieren, sowohl auf Jungen im Hinblick auf die öffentliche Darstellung als „generelle“ Schulversager als auch auf jungenuntypische Leistungsdomänen wie Lesen (Plante, Théorét & Favreau, 2009) übertragen werden.

Durch die präsentierten Befunde zum Zusammenspiel von *stereotype threat*, Motivation und Leistung kann auch von einer Verschärfung der Problematik ausgegangen werden, da neben den bereits beschriebenen Auswirkungen auf Schulleistungen und Noten nun noch hinzukommt, dass die Stereotypenbedrohung für eine Verfestigung der Leistungsdifferenzen sorgt. Anders gesprochen kann abgeleitet werden, dass Stereotypenbedrohung einerseits zu schlechteren Leistungen führt, andererseits aber *zusätzlich* dafür sorgt, dass Intentionen zur individuellen Leistungsverbesserung und damit zur Beseitigung von Kompetenz- und Leistungsunterschieden unwahrscheinlicher werden (vgl. Fogliati & Bussey, 2013), da die Lernmotivation zur Auseinandersetzung mit einem stereotypisierten Lerngegenstand einer ungünstigen Beeinflussung durch *stereotype threat* unterliegt. Dies kann zur Erklärung beitragen, wie sich relativ kurzfristige Auswirkungen einer Stereotypenbedrohung dauerhafter implementieren und zur Beschreibung längerfristiger Unterschiede hinsichtlich des Bildungserfolgs von Jungen beitragen können.

4. Das Soziale als Teil des Selbst: Die Theorie der sozialen Identität

Nach der einführenden Darstellung über den Bildungserfolg von Jungen und der darauf aufbauenden, medialen Darstellung sowie von theoretischen Annahmen und empirischen Befunden zu möglichen Auswirkungen im Sinne einer Stereotypenbedrohung für Jungen wird im Folgenden darauf eingegangen, welche Reaktionen von Jungen auf die Bedrohung als „Schulversager“ erwartet werden können. Die *Social Identity Theory* von Tajfel und Turner (1986) erlaubt hierzu die Ableitung von vielfältigen Strategien, um die positive Distinktheit der *Ingroup* der Jungen von der *Outgroup* der Mädchen wiederherzustellen. Auf diese soll im nächsten Abschnitt nach einer einführenden Erläuterung der theoretischen Annahmen von Tajfel & Turner (1986) genauer eingegangen werden, um herzuleiten ob und wie auf eine Bedrohungssituation auch über die Effekte auf Leistung und Motivation im Sinne der *stereotype threat*-Theorie hinaus Bewältigungsstrategien eingesetzt werden können, um die potentiell bedrohte männliche Identität wiederherzustellen und zu stabilisieren.

4.1. Entstehungszusammenhang

In der Sozialpsychologie der 1980er Jahre wurde bemängelt, dass das Wissenschaftsgebiet *per definitionem* untersucht, wie sich Verhalten und Erleben von Menschen in Gegenwart anderer Individuen gestaltet und welchen Einfluss andere Individuen auf eigenes Verhalten und Erleben haben (zur Entstehungsgeschichte vgl. Hogg & Williams, 2000; Turner & Reynolds, 2010; siehe auch: Allport, 1968). Konträr zu dieser Definition beschäftigten sich damalige Forschungsarbeiten allerdings in der Hauptsache mit individuellen Prozessen und verlagerten somit den Forschungsfokus weg von der sozialen hin zur individuellen Kognition. Unter anderem kritisierte Tajfel (1972) diese Entsozialisierung der Sozialpsychologie und forderte einen stärkeren Einbezug sozialer Strukturen und sozialer Interaktion, um sich der ursprünglichen Definition des Wissenschaftsbereichs wieder anzunähern. Dem Individuum kommt zwar als eigentlich handelnder Akteur nichtsdestotrotz eine bedeutsame Rolle zu, jedoch ist es in der sozialen Realität häufig anzutreffen, dass Individuen ein bestimmtes Verhalten zeigen, diese sich aber als Repräsentant/-in einer sozialen Gruppe wahrnehmen und dementsprechend ihre Handlungen und Absichten eher daran und weniger individuell ausrichten (Mummendey, 1985). Es wird somit deutlich, dass Theorien, die individuelles Verhalten erklären nicht ohne weiteres auf den sozialen Kontext übertragen werden können, sondern vielmehr gesonderte theoretische Ansätze erforderlich sind.

Die Forschungsarbeiten von Sherif (für einen Überblick: Jackson, 1993; Sherif, 1967) greifen die Notwendigkeit unmittelbar auf, dass soziales Verhalten im entsprechenden Kontext untersucht werden muss und liefern somit eine erste theoretische Rahmung von Intergruppenver-

halten. Nach Annahmen von Sherif (1967) ist es psychologisch hochgradig bedeutsam und mit entsprechenden Konsequenzen verbunden, sich als Gruppenmitglied wahrzunehmen und sich entsprechend zu verhalten. Die von der Arbeitsgruppe um Sherif entwickelte Theorie des realistischen Gruppenkonflikts sollte den o.g. Ansprüchen an eine sozialpsychologische Erklärung von individuellem Verhalten in einer sozialen Gruppe gerecht werden. Die Theorie sagt vorher, dass die Herausbildung von Vorurteilen und Feindseligkeiten sowie die Genese von diskriminierendem Verhalten gegenüber Gruppenmitgliedern einer *Outgroup* i.d.H. durch einen Interessenskonflikt zwischen den verschiedenen Gruppen verursacht wird. Es werden somit die Interessen und Ziele zweier, um knappe Ressourcen konkurrierender Gruppen als Grundlage beschrieben, welche eine entscheidende Funktion für die Ausrichtung des Verhaltens einnehmen. Hieraus resultiert, dass beide Gruppen nicht unabhängig voneinander sind – in dieser wechselseitigen Abhängigkeit (Interdependenz) liegt ein wichtiger Schlüssel zum Verständnis von Intergruppenverhalten (Bornstein, 2003). Wenn negative Interdependenz zwischen Gruppen besteht, geht die Zielerreichung für eine Gruppe mit der Entfernung vom Ziel der anderen Gruppe einher (v.v. bei positiver Interdependenz). Praktisch angewendet bedeutet dies, dass Einstellungen und Verhalten zwischen Gruppen abhängig von real existierenden Konflikten zwischen den Gruppen sind. Im Rahmen verschiedener experimenteller Studien wurden diese theoretischen Annahmen überprüft (Sherif, Harvey, White, Hood & Sherif, 1961). In vier aufeinander aufbauenden Phasen wurde modelliert, wie es zur Entstehung und zur Lösung von Gruppenkonflikten kommt. Hierzu wurden Jungen während eines Ferienlagers ungeachtet ihrer bestehenden Freundschaften zunächst in zwei Gruppen aufgeteilt, die darauf folgend innerhalb gemeinsamer Interaktionen eine stabile Gruppenstruktur herausbildeten. Es folgten verschiedene Wettkämpfe, in welchen die Konkurrenz zwischen beiden Gruppen herausgestellt wurde. In diesem Abschnitt des Experiments wurde beobachtet, dass Diskriminierung einhergehend mit abwertendem, feindseligem Verhalten gegenüber der *Outgroup* verstärkt auftrat. In einer letzten Phase der Untersuchung wurden beide Gruppen mit einem Problem konfrontiert, welches nur gemeinsam gelöst werden kann (*positive Interdependenz*). Im Rahmen der kooperativen Problemlösung beider Gruppen verringerte sich das zuvor verstärkt aufgetretene, feindselige Verhalten zwischen den Gruppen zunehmend.

Im Sinne der Theorie des realistischen Gruppenkonflikts wurden diese Ergebnisse dahingehend interpretiert, dass einerseits bestätigt werden konnte, dass Interessenkonflikte zwischen Gruppen eine Ursache für diskriminierendes Verhalten gegenüber einer Außengruppe darstellen, wobei als handelnde Akteure Einzelpersonen durch ihre Mitgliedschaft in einer sozialen Gruppe in ihren Handlungen geleitet werden. Ferner wurde deutlich, dass bei positiver Abhängigkeit, d.h. wenn ein übergeordnetes Ziel nur durch Zusammenarbeit mit der *Outgroup* erreicht werden kann, entsprechende kooperative soziale Interaktionen vorherrschen und freundliche,

die Gruppen verbindende Einstellungen entstehen. Zusätzlich wurde festgestellt, dass antagonistische Beziehungen zu einer Außengruppe funktional für die *Ingroup* im Sinne einer erhöhten Verbundenheit innerhalb der Gruppe und einer verstärkten Bevorzugung der Eigengruppe erweist. Somit kommt der Konkurrenz um knappe Ressourcen und der daraus entstehenden Funktion der Intergruppenbeziehung für das Erreichen eines Ziels der Theorie folgend eine zentrale Bedeutung für das Verständnis von Intergruppenverhalten zu (Jackson, 1993). Kritisch wurde angemerkt, dass ohne Kontrollgruppe nicht festzustellen ist, inwiefern der Interessenskonflikt und daraus entstehender sozialer Wettbewerb zwischen Gruppen eine notwendige Bedingung für diskriminierendes Intergruppenverhalten darstellt (van Knippenberg, 1978) oder ob bereits weniger starke Konkurrenz um ein bestimmtes Ziel eine Eigengruppenbevorzugung und eine Diskriminierung der *Outgroup* zur Folge hat (Ferguson & Kelley, 1964; Rabbie & Wilkens, 1971).

Um dieser Fragestellung nachzugehen, wurde mit verschiedenen experimentellen Arbeiten, die zusammenfassend als *minimal group*-Paradigma bezeichnet werden, der Grundstein für die spätere Entwicklung der Theorie der sozialen Identität (Tajfel & Turner, 1986) gelegt. In mehreren Experimenten (Tajfel, 1970; Tajfel, Billig, Bundy & Flament, 1971) wurden Versuchspersonen, zufällig in zwei Gruppen eingeteilt. Als Gemeinsamkeit der Gruppe wurde eine fiktive Rückmeldung über die angebliche Bevorzugung eines Künstlers (Klee-Gruppe vs. Kandinsky-Gruppe) vorgelegt. Es fand zu keinem Zeitpunkt irgendeine Form sozialer Interaktion statt und es wurde auch keine ebensolche in Aussicht gestellt. Hernach wurden die Probanden/-innen gebeten, fiktive Geldbeträge oder geldwerte Punkte an zwei andere Versuchsteilnehmer/-innen zu verteilen (eine Verteilung an sich selbst war ausgeschlossen). Es gab keinerlei Informationen, über die Identität der Personen oder sonstige Randfaktoren, lediglich die Gruppenmitgliedschaft wurde salient gemacht, d.h. es wurde mitgeteilt, dass eine bestimmte Person zur Eigen- und die andere Person zur Fremdgruppe gehört. Da es wie beschrieben keinen sozialen Kontakt zu Mitgliedern der Eigen- oder Fremdgruppe gab, erfolgte die Konstruktion der Gruppensituation ausschließlich durch rein kognitive Leistungen „im Kopf“. *In absentio* einer realen Situation wurde dennoch eine solche konstruiert und durch die Probanden/-innen als real existierende Gruppe wahrgenommen, was nunmehr auch Tajfel & Turner (1986) im Rahmen ihrer Theorie als *minimal group* bezeichnen.

Mummendey (1985) fasst die grundlegenden Voraussetzungen für das *minimal group*-Paradigma wie folgt zusammen:

- (1) keinerlei face-to-face Interaktionen mit Eigen- oder Fremdgruppenmitgliedern
- (2) Anonymität der Gruppenmitglieder: Außer der Zugehörigkeit zur Gruppe gibt es keinerlei Informationen über die beteiligten Personen

- (3) Gruppeneinteilung nach trivialen *ad hoc* Kriterien / Anm.: Selbst bei offensichtlich zufälliger Gruppenzuweisung lassen sich u.g. Effekte nachweisen (Billig & Tajfel, 1973)
- (4) Verhalten der Versuchsperson kann keinen persönlichen Nutzen bringen
- (5) gezeigte Verhaltensweisen müssen für die Versuchsperson real und von Bedeutung sein (konkrete Belohnungen bzw. Bestrafungen)

Die Ergebnisse dieser Untersuchung zeigten, dass die Teilnehmenden am Versuch der Eigengruppe systematisch mehr Ressourcen (Geld, Punkte) zuwies, als der *Outgroup* und die obwohl die Zusammenstellung der Gruppe auf irrelevanten Merkmalen wie der angeblichen Präferenz eines Künstlers erfolgte (vgl. auch: Brewer, 1979; Mullen, Brown & Smith, 1992). Ergänzend zeigte sich deutlich, dass die Bevorzugung der Eigengruppe nicht so erfolgte, dass diese maximal profitiert, sondern dass die Unterschiede zwischen den der Eigen- und Fremdgruppe zugewiesenen Beträge maximal wurde (*maximum difference*), auch wenn hierbei nicht der maximale Gewinn für die Eigengruppe erzielt werden konnte. Der Theorie des realistischen Gruppenkonflikts nach Sherif et al. (1961) setzt als notwendige Bedingung für den Intergruppenkonflikt eine reale Konkurrenz um knappe Ressourcen voraus, was zur Folge hätte, dass Probanden/-innen der *minimal group*-Experimente der Eigengruppe stets den maximal möglichen Betrag zuweisen und nicht nach maximaler Differenz zur *Outgroup* streben. Um diese Abweichung erklären zu können, wurde aufbauend auf das *minimal group*-Paradigma die Theorie der sozialen Identität (Tajfel & Turner, 1986) formuliert und eine schlüssige Erklärung der Befunde aus den *minimal group*-Experimenten abgeleitet.

4.2. Grundannahmen der Social Identity Theory: Theoretische Kontinua

Tajfel und Turner (1986) definieren eine soziale Gruppe als eine Mehrzahl von Menschen, die sich als Mitglied einer bestimmten sozialen Kategorie wahrnehmen und mit der Gruppenmitgliedschaft bestimmte emotionale Qualitäten verbinden sowie damit einhergehend die eigene Gruppe wie auch die eigene Mitgliedschaft in der Gruppe bewerten. Der Wahrnehmung des Individuums als Mitglied einer sozialen Gruppe kommt demnach eine besondere Bedeutung zu. Gleichwohl ist es möglich und aufgrund der zahlreichen denkbaren, kategorisierenden Variablen auch hochgradig wahrscheinlich, dass sich eine Person als Mitglied mehrerer sozialer Gruppen wahrnimmt (z.B. bezüglich der Geschlechterzugehörigkeit, der Professionalisierung, des Lebensalters usw.). Mit vier korrespondierenden theoretischen Kontinua (Darstellung nach Mummendey, 1985) soll versucht werden, im Sinne der *Social Identity Theory* zu erklären, wann sich Personen als Individuum wahrnehmen und entsprechendes Verhalten zeigen bzw. wann die Wahrnehmung als Mitglied einer sozialen Gruppe dominiert und darauf aufbauend das eigene Handeln auf Ebene einer sozialen Gruppe ausgerichtet und entsprechend beeinflusst wird. An

dieser Stelle sei bereits angemerkt, dass Tajfel und Turner (1986) darauf hinweisen, dass nur sehr wenige soziale Situationen denkbar scheinen, bei denen sich soziales Verhalten eindeutig einem Pol der theoretischen Kontinua zuweisen lässt, sondern im Regelfall ein breites Spektrum auf den theoretischen Kontinua ausgenutzt wird, wenn soziales Verhalten analysiert und klassifiziert werden soll.

Das erste Kontinuum differenziert danach, ob Verhalten eindeutig interpersonal bzw. eindeutig intergruppal ausgerichtet ist. Situationen nahe des interpersonalen Pols zeichnen sich dadurch aus, dass interaktives Verhalten zwischen zwei oder mehr Personen vollkommen unbeeinflusst von sozialen Kategorien ist, denen die Interaktionspartner/-innen angehören. Vielmehr ist eine derartige Interaktion ausschließlich durch individuelle Eigenschaften der interagierenden Person gekennzeichnet. Im Gegenzug dazu wird nahe des intergruppalen Extrems die Zugehörigkeit zu einer sozialen Gruppe bzw. das Vorhandensein verschiedener sozialer Kategorien vollständig verhaltens- und handlungsleitend, individuelle Eigenschaften treten vollkommen in den Hintergrund.

Das zweite theoretische Kontinuum bezieht sich auf Einstellungen, Werte und Überzeugungen über die Beziehungen sozialer Gruppen sowie deren Veränderbarkeit und ist durch die Extrema „soziale Mobilität“ vs. „soziale Veränderung“ gekennzeichnet. Während bei sozialer Mobilität die Überzeugung vorherrscht, dass innerhalb eines sozialen Systems ein Wechsel von einer Gruppe zu einer anderen Gruppe leicht möglich ist, geht der Gegenpol „soziale Veränderung“ davon aus, dass es schwierig bis unmöglich ist, die eigene Gruppe zu verlassen und damit einhergehend ein Wechsel der Gruppenmitgliedschaften (wenn überhaupt) nur im Rahmen eines Wandels eines gesamten sozialen Systems möglich wäre. Denkbar wäre diesbezüglich beispielsweise die Wahrnehmung auf Grundlage der eigenen (biologischen) Geschlechtszugehörigkeit, welche die Mitgliedschaft innerhalb der sozialen Gruppe „Männer“ oder „Frauen“ determiniert und für zahlreiche Mitglieder dieser sozialen Gruppe eine Änderung der Gruppenzugehörigkeit unmöglich scheint.

Als drittes theoretisches Kontinuum wird postuliert, dass soziale Situationen auf Gruppenebene danach zu unterscheiden sind, inwiefern sie eine individuelle Variabilität zulassen bzw. berücksichtigen. Gemeint ist damit, dass am ersten Extrempunkt das Verhalten innerhalb einer *Ingroup* eine große Varianz aufweist, während am anderen Pol ein Maximum an „Gleichförmigkeit“ (Mummendey, 1985) innerhalb der Gruppe erreicht wird. An diesem theoretischen Kontinuum wird die Verknüpfung aller theoretischen Kontinua untereinander besonders deutlich, da eine sehr geringe individuelle Variabilität innerhalb der Gruppe häufig mit Situationen einhergeht, welche auf dem ersten Kontinuum nahe des intergruppalen Pols liegen.

Im vierten und letzten Kontinuum wird die Beziehung von Gruppenmitgliedern zwischen *In-* und *Outgroup* in den Fokus gerückt. Am ersten Extrempunkt dieses Kontinuums ist die Wahrnehmung, Beurteilung und Behandlung von Mitgliedern der *Outgroup* durch die individuell unterschiedlichen Merkmale der *Outgroup* gekennzeichnet. Hier wird also verstärkt einbezogen, dass eine Außengruppe intragrupal durch eine individuelle Verschiedenheit geprägt ist. Konträr dazu geht das entgegengesetzte Extrema von einer maximalen Homogenität innerhalb der *Outgroup* aus, was Tajfel (1978) mit Wahrnehmung von Mitgliedern der Außengruppe als „undifferenzierte Items“ beschreibt. Erneut wird deutlich, wie die Kontinua untereinander korrespondieren: Je eher eine soziale Situation auf dem ersten Kontinuum nahe des interpersonalen Extremas und auf dem zweiten Kontinuum nahe der „sozialen Mobilität“ angesiedelt ist, desto wahrscheinlicher ist auf dem vierten Kontinuum eine Wahrnehmung der Außengruppe und ihrer Mitglieder als individuell verschieden, was entsprechende Konsequenzen für das eigene Verhalten hat. Im Gegensatz dazu führt eine Wahrnehmung sozialer Situationen an den Gegenpolen (intergruppal und „soziale Veränderung“) eher zu einer Homogenisierung der *Outgroup*-Mitglieder als „undifferenzierte Items“ (Mummendey, 1985; Tajfel & Turner, 1986).

4.3. Verhaltenskonsequenzen auf Intergruppenebene

Die Theorie der sozialen Identität basiert auf drei zentralen Grundannahmen (nach Turner, 1999), und zwar dass (a) Individuen danach streben, eine positive Selbsteinschätzung zu erhalten oder aber eine möglicherweise negative Selbsteinschätzung zu verbessern, (b) ein bestimmter Teil dieser Selbsteinschätzung ist die soziale Identität, die sich aus den sozialen Gruppenmitgliedschaften und den Evaluationen dieser Mitgliedschaften ergibt sowie (c) die Evaluation der Gruppenmitgliedschaft ergibt sich aus Vergleichen der Eigengruppe mit signifikanten Außengruppen (vgl. auch: Jetten, Spears & Manstead, 1999). Bereits Festinger (1954) wies in seiner sozialen Vergleichstheorie darauf hin, dass eigene Gefühle und Einschätzungen mit denen signifikanter anderer Personen verglichen werden, um die Angemessenheit der eigenen Gefühle und Einschätzungen abgleichen zu können. Diese Annahme wird in der Theorie der sozialen Identität aufgegriffen und erweitert: Positive Vergleiche steigern, negative Vergleiche senken das wahrgenommene Prestige der Eigengruppe. Dieses Streben nach einem besseren Abschneiden der Eigengruppe bei Vergleichen wird als Streben nach positiver Distinktheit der *Ingroup* bezeichnet. Bedingungen für das Funktionieren der vorgenannten Prinzipien stellen eine Internalisierung der eigenen Gruppenmitgliedschaft, eine Bedeutsamkeit der Gruppenmitgliedschaft in einer bestimmten sozialen Situation und der Vergleich mit einer relevanten, bedeutsamen Außengruppe und nicht irgendeiner beliebigen *Outgroup* dar (vgl. auch: Tajfel & Turner, 1986).

Kern der Theorie stellen somit vier miteinander verbundene Konzepte dar, nämlich soziale Kategorisierung, soziale Identität, sozialer Vergleich und soziale Distinktheit. Im Rahmen der sozialen Kategorisierung wird die Umwelt in verschiedene Teilbereiche zergliedert und anhand spezifischer, sozialer Merkmalsdimensionen in voneinander unterscheidbare soziale Kategorien eingeordnet. Durch diese Kategorisierung wird erst ermöglicht, eine klare Gruppenzugehörigkeit – z.B. auf Basis von Geschlecht, Einkommen oder Alter – zu etablieren. Neben einer Orientierungsfunktion, wer zur Eigen- und wer zur Fremdgruppe gehört, bietet dieser Prozess auch eine Positionsbestimmung der eigenen Person innerhalb einer komplexen Realität (Mummendey, 1985). Aus den so entstehenden Mitgliedschaften in sozialen Gruppen speist sich die soziale Identität. Diese kann als der Teil des Selbstkonzeptes angesehen werden, der aus dem Wissen um seine eigene Zugehörigkeit zu bestimmten sozialen Gruppen besteht und eine emotionale Bewertungskomponente dieser Gruppenzugehörigkeit beinhaltet (Tajfel, 1982).

Um überhaupt zu einer Bewertung von Gruppenmitgliedschaften zu kommen, ist es notwendig die eigene Gruppe durch möglichst vielschichtige soziale Vergleiche zu evaluieren. Für den Vergleich mit einer *Outgroup* gilt es allerdings, innerhalb eines sozialen Wettbewerbs relevante Vergleichsdimensionen und eine relevante (i.S.v. auf einigen Merkmalsdimensionen ähnliche) Vergleichsgruppe heranzuziehen. Ebenso bedeutsam ist die Wahrnehmung der Stabilität sowie der Legitimität von Unähnlichkeiten zwischen Gruppen für einen möglichst aussagekräftigen Vergleich zwischen Eigen- und Fremdgruppe.

Schlussendlich führt das Bedürfnis nach einer Unterschiedlichkeit zwischen *In-* und *Outgroup* einhergehend mit der bereits zuvor getroffenen theoretischen Annahme, dass jedes Individuum nach einer positiven Selbstbewertung strebt, dazu, dass im letzten Schritt nach der Theorie eine positive Distinktheit der Eigengruppe angestrebt wird (Haslam, Ellemers, Reicher, Reynolds & Schmitt, 2010). Im Falle von negativ ausfallenden Gruppenvergleichen oder einer wahrgenommenen Bedrohung der Überlegenheit der Eigengruppe lassen sich aus den theoretischen Annahmen, insbesondere den Annahmen über die Vergleichsprozesse zwischen Gruppen, Möglichkeiten ableiten, mit einer bedrohten Statusüberlegenheit oder einer festgestellten Statusunterlegenheit umzugehen, um wiederum eine positive Distinktheit für die Eigengruppe herzustellen und die bedrohte soziale Identität zu stabilisieren (vgl. auch Branscombe, Ellemers, Spears & Doosje, 1999).

Zum einen kann entsprechend des zweiten theoretischen Kontinuums mehr oder minder einfach die betroffene Gruppe verlassen werden, um Teil einer neuen (statushöheren) zu werden. Dieser soziale Aufstieg setzt jedoch voraus, dass die entsprechenden gruppendifinierenden Merkmale veränderbar sind und der Zielgruppe entsprechend angepasst werden können. Weiterhin bietet sich die Möglichkeit, direkten sozialen Wettbewerb zu betreiben. Das bedeutet,

dass bei einem positiven Vergleichsergebnis die positive Distinktheit stark überbetont und bewusst inszeniert wird – es werden also vermehrt Verhaltensweisen gezeigt, welche die vermeintliche Überlegenheit der Eigengruppe darstellt. Ferner ließe sich auf Basis des direkten sozialen Wettbewerbs die Abwertung der *Outgroup* als Strategie ableiten, d.h. es würden in diesem Falle vermehrt Verhaltensweisen gezeigt, die darauf abzielen den Status dieser Gruppe abzuwerten und damit eigene (wahrgenommene) Überlegenheit zu stabilisieren und kommunizieren. Wie ersichtlich wird, eignen sich Strategien auf Grundlage des direkten sozialen Wettbewerbs eher für Gruppen, die eine Statusüberlegenheit wahrnehmen und werden als *Eigengruppenfavorisierung* bezeichnet. Bei wahrgenommener Statusunterlegenheit wären diese Verhaltensweisen insbesondere unter Rückbezug auf die theoretische Anforderung eines *realistischen* Gruppenvergleichs (mit Ausnahme der Abwertung einer Fremdgruppe) weniger zielführend. Ergänzend werden daher weitere Möglichkeiten unterschieden, welche in Anlehnung an Tajfel & Turner (1979) als „soziale Kreativität“ bezeichnet werden. So kann eine bedrohte Gruppe neue Vergleichsdimensionen entwerfen, auf der sie wiederum besser abschneidet, die ursprüngliche Bewertung umkehren, sodass der eigentlich ungünstigere Vergleich nunmehr als besonders positiv und wünschenswert interpretiert wird oder die Vergleichsgruppe ändern, sich also z.B. mit vergleichsweise statusniedrigeren Gruppen vergleichen. Zu berücksichtigen bleibt bei diesen Strategien, dass insbesondere der Entwurf neuer Vergleichsdimensionen und die Wahl einer anderen Vergleichsgruppe nur in einem eingeschränkten Maße stattfinden können, da zu unrealistische Wahlen neuer Vergleichsdimensionen und -gruppen die theoretischen Anforderungen nach einem Intergruppenvergleich mit *signifikanten* Vergleichsdimensionen und *realistischen* Vergleichsgruppen verletzen. Dies wiederum hätte der Theorie nach zur Folge, dass keine positive Distinktheit erzeugt werden könnte und somit die bedrohte soziale Identität nicht stabilisiert werden kann. Abschließend bleibt festzuhalten, dass die Strategien der „sozialen Kreativität“ ebenfalls nur wirksame Instrumente zur Herstellung positiver Distinktheit darstellen, wenn die Resultate der Uminterpretation des sozialen Vergleichs sowohl von der gesamten *Ingroup* anerkannt als auch von der *Outgroup* angenommen werden (Tajfel & Turner, 1986; Turner & Reynolds, 2001).

Mit diesen theoretischen Annahmen lassen sich nun auch die eingangs dargestellten Ergebnisse der *minimal group*-Paradigmen einordnen. Nach der Theorie des realistischen Gruppenkonflikts von Sherif et al. (1961) wäre zu erwarten, dass die beiden konkurrierenden Gruppen versuchen, einen maximalen Gewinn für die Eigengruppe zu erzielen. Tatsächlich aber wurde festgestellt, dass die Probanden/-innen der *minimal group*-Experimente vielmehr eine Strategie verfolgten, nach welcher die Differenz zwischen den Gruppen maximal wird („*maximum difference*“; Tajfel, 1970). Unter Anwendung der Theorie der sozialen Identität lassen sich diese

Ergebnisse nunmehr schlüssiger erklären: Individuelle Mobilität ist im Fall der *minimal group*-Experimente nicht möglich, denn durch die Einteilung in zwei Gruppen bestand keine Möglichkeit, die Gruppe zu verlassen und einer anderen Gruppe beizutreten. Strategien der sozialen Kreativität zur Herstellung einer positiven Distinktheit zur Außengruppe bieten ebenso keine ausreichende Alternativerklärung, da die Versuchspersonen außer dem zu verteilenden Geld über keinerlei alternative Dimension verfügten, auf der ein Vergleich stattfinden konnte. So bleiben Strategien des sozialen Wettbewerbs, um das Verhalten der Probanden/-innen zu erklären: Da das Geld nicht im Sinne einer Wertdimension oder eines einzuholenden Gewinns wirkte, sondern einzig und allein eine Dimension des sozialen Wettbewerbs darstellte, wählten die Probanden Verteilmechanismen aus, welche die *Ingroup* zwar bevorzugen, aber nicht den maximalen Gewinn erzielen. Vielmehr ist von zentraler Bedeutung geworden, dass die Verteilung der monetären Ressourcen eine maximale Unterschiedlichkeit zwischen In- und *Outgroup* herstellt. Dies wiederum stellt sicher, dass sich die Eigengruppe maximal positiv (auf der einzig verfügbaren Vergleichsdimension) von der Außengruppe abhebt und somit ein Höchstmaß an positiver Distinktheit hergestellt wird, was schlussendlich eine positive soziale Identität erzeugt und absichert.

Der folgende Abschnitt soll einen kurzen Überblick über die bisherige Forschung geben, welche auf den Annahmen der Theorie der sozialen Identität beruht und geprüft hat, ob sich die theoretischen Annahmen auch empirisch validieren lassen.

4.4. Bedeutung und empirische Überprüfung der Social Identity Theory

Turner (1999) fasst in einem Rückblick zusammen, dass die Theorie der sozialen Identität vollkommen neue Sichtweisen auf Gruppen, Intergruppenprozesse, Vorurteile und Stereotypisierung sowie auf das Selbst und die Persönlichkeit erlaubt. Eine empirisch zahlreich überprüfte Grundannahme ist die Tatsache, dass in sozialen Kontexten eine Depersonalisation stattfindet, d.h. eine geteilte, soziale Identität salient wird und damit einhergehend die Wahrnehmung der eigenen Person und anderer Individuen entscheidend von individuellen Prozessen losgelöst wird (Simon, Pantelo & Mummendey, 1995; Spears, Doosje & Ellemers, 1997). Man agiert also nicht mehr als einzelne Person, sondern denkt und handelt im Sinne seiner derzeit salienten, sozialen Gruppe, was Turner (1999) als „*shift from action in terms of differing personal identities to action more in terms of a shared social identity*“ (S. 15) bezeichnet.

In Bezug auf die Stereotypisierung wurde in der frühen, sozialpsychologischen Forschung angenommen, dass es sich hierbei um einfache kognitive Vereinfachungen und Verzerrungen handelt, welche für die Bewältigung von kognitiv komplexen Situationen funktional sind, d.h. dabei helfen, Komplexität im Sinne einer kognitiven Ökonomie zu reduzieren (Turner, 1999).

Die Theorie der sozialen Identität erlaubt eine vollkommen neue Betrachtungsmöglichkeit und argumentiert, dass Stereotype aus sozialer Wahrnehmung von Gruppen und ihren Mitgliedern entstehen und für Intergruppenprozesse funktional zur Sicherung positiver Distinktheit sind (Oakes, Haslam & Turner, 1994). Stereotype stellen somit Kategorisierungen auf Ebene der sozialen Identität dar, in welcher Individuen in Form von Eigenschaften und Charakteristika der sozialen Gruppe, der sie angehören (d.h. auf Basis von Inter- oder Intragruppenbeziehungen), definiert werden. Die Variabilität und Kontextabhängigkeit wurde in verschiedenen empirischen Arbeiten nachgewiesen und es konnte gezeigt werden, dass ein Stereotyp ein und derselben Person in Art und Inhalt je nach sozialem Kontext und Intergruppenbeziehung variieren kann und nicht wie bisher angenommen eine fixe, unveränderliche Variable darstellt (Haslam & Turner, 1992; Haslam, Turner, Oakes, McGarty & Hayes, 1992; Oakes et al., 1994).

Die Theorie der sozialen Identität erweitert zusätzlich das bis dato vorherrschende Verständnis von Intergruppenprozessen und geht davon aus, dass Verhalten zwischen Gruppen hauptsächlich durch ein Bedürfnis nach maximaler Differenzierung der Eigen- von der Fremdgruppe und dem Streben nach positiven Vergleichsergebnissen mit relevanten Fremdgruppen zur Absicherung einer positiven Distinktheit der *Ingroup* determiniert ist. Empirische Unterstützung für die durch die *Social Identity Theory* postulierten Intergruppenprozesse im Sinne eines Strebens nach maximaler Differenzierung und positiver Distinktheit legen Jackson, Sullivan, Harnish & Hodge (1996) vor. In drei Experimenten prüften sie sowohl die Gültigkeit der Annahme zur sozialen Mobilität als auch zur sozialen Kreativität. Zusätzlich wurden die Randbedingungen entsprechend der theoretischen Kontinua, insbesondere zur Durchlässigkeit der Gruppengrenzen (soziale Mobilität vs. sozialer Wandel), als ergänzende Variable in den Experimenten geprüft.

In drei Experimenten – einem *minimal group*-Paradigma und zwei realen Gruppensituationen – findet sich empirische Evidenz, dass (1) negativ distinkte Mitglieder einer *Ingroup* sich als weniger prototypisch für die Eigengruppe wahrnehmen (Vergleichsdimensionen ändern), (2) Mitglieder der *Ingroup* das Kriterium, auf welchem die negative Distinktheit beruht, insgesamt nicht so negativ wahrnehmen wie Mitglieder einer Kontrollgruppe (Änderung der Valenz des bisher negativen Vergleichsmerkmals), (3) Mitglieder einer negativ distinkten *Ingroup* auf anderen Vergleichsdimensionen die Eigengruppe favorisieren und (4) bei undurchlässigen Gruppengrenzen die Eigengruppenfavorisierung und die Uminterpretation der negativen Vergleichsdimension in eine positive verstärkt wird (Jackson et al., 1996). Für die Annahme, dass (5) bei durchlässigen Gruppengrenzen Mitglieder eher zu einem Wechsel der Gruppe tendieren und sich daher verstärkt als nicht prototypisch für die Eigengruppe wahrnehmen konnte keinerlei Unterstützung gefunden werden.

Insgesamt lässt sich somit festhalten, dass für vier der fünf überprüften Mechanismen der Theorie der sozialen Identität empirische Evidenz gefunden werden konnte, welche die Gültigkeit der Annahmen der *Social Identity Theory* unterstützen. Es konnte gezeigt werden, dass bei negativer Distinktheit die Eigengruppe sukzessive aufgewertet wird und auch soziale Mobilität stattfindet, indem sich ein Individuum je nach Ergebnis des Intergruppenvergleichs als prototypisch oder weniger prototypisch für die Gruppe wahrnimmt. Im Falle eines negativen Vergleichs bietet die Wahrnehmung des Selbst als nicht prototypisches Gruppenmitglied die Möglichkeit, die soziale Bedrohung in gewissem Maße abzuschwächen und durch soziale Kreativität den Versuch zu unternehmen, eine positive Distinktheit zu erreichen. Bei einem positiven Intergruppenvergleich dient hingegen die Wahrnehmung der eigenen Person als prototypisch für die Gruppe der Herstellung und Absicherung einer positiven sozialen Identität (vgl. auch: Ellemers, 1993; Ellemers, Wilke & van Knippenberg, 1993; Turner & Brown, 1978).

In einer Metaanalyse zum *ingroup bias* prüften Mullen et al. (1992) insgesamt 137 Studien und hielten zusammenfassend fest, dass die Eigengruppenbevorzugung in allen in die Analyse eingegangenen Studien nachgewiesen werden konnte. Diese Grundannahme der Theorie der sozialen Identität kann somit als gesichert betrachtet werden. Allerdings werden Kontextfaktoren identifiziert, welche die Stärke des Effekts nachhaltig beeinflussen. Wenn eine sehr saliente Gruppensituation geschaffen wurde, fiel der *ingroup bias* vergleichsweise höher aus. Die Realität bzw. Salienz der Gruppenkategorisierung interagiert weiterhin bedeutsam mit dem Status bzw. dem Prestige einer Gruppe: Es zeigte sich ein leichter Rückgang des *ingroup bias* in Abhängigkeit des Status in realen Gruppen und eine deutliche Steigerung des *ingroup bias* in Abhängigkeit des Gruppenstatus in „künstlichen“ Gruppen. Ferner fällt bei Gruppen mit hohem Prestige die Verzerrung zugunsten der Eigengruppe besonders hoch auf für die Gruppe hochgradig relevanten Vergleichsdimensionen aus, wohingegen bei Gruppen mit vergleichsweise weniger Prestige eine Verzerrung nur auf peripheren Vergleichsdimensionen auftritt. Diese Ergebnisse liefern zum einen wichtige Evidenz für eine Kernannahme der *Social Identity Theory*, weisen zum anderen aber darauf hin, dass Randbedingungen wie die Realitätsnähe der Gruppenkategorisierung und des wahrgenommenen Status der Gruppe berücksichtigt werden müssen. Einen weiteren Hinweis über die Wirksamkeit der Eigengruppenverzerrung liefern Oakes & Turner (1980) sowie Lemyre & Smith (1985), die zeigten, dass sich Probanden/-innen nach derartigen Verzerrungen – im Sinne einer stabilisierten sozialen Identität – besser fühlen. Dies macht empirisch deutlich, was die Theorie der sozialen Identität vorhersagt: Differenzierung zwischen *In-* und *Outgroup* und damit einhergehend die Favorisierung der Eigengruppe resp. die Diskriminierung der Außengruppe findet statt, damit Individuen sich als *Gruppe*, aber auch *individuell* auf Ebene der eigenen sozialen Identität wohlfühlen. Auf den Altersbereich der Jugend angewandt demons-

triert Tarrant (2002) exemplarisch, dass der Mechanismus der Eigengruppenfavorisierung auch im Jugendalter angewendet wird und eine identische Funktionalität aufweist.

Abschließend eröffnet die Theorie einen neuen Blickwinkel auf Aspekte von Selbst und Persönlichkeit. Diese Bereiche, die in der bisherigen psychologischen Forschung als hochgradig individuell und persönlich betrachtet wurden, werden durch die Theorie der sozialen Identität auf einen sozialen Kontext bezogen. Während bisher individuell kognitionspsychologische Erklärungen für die Entstehung des Selbst herangezogen wurden und eine Annahme über das Selbst als starre, separate mentale Struktur dominierte (Turner, 1999), macht die Theorie der sozialen Identität deutlich, dass Selbst und Persönlichkeit in hohem Maße von sozialen Bedingungen und damit einhergehend auch Gruppenprozessen mitbestimmt werden und dass z.B. Selbstwahrnehmungen in verschiedenen sozialen Kontexten als dynamisches Konstrukt aufzufassen sind (Turner & Oakes, 1997). Das Bild des Selbst wird damit aus der Betrachtung interpersonaler, individueller Kognition herausgerückt und „sozialisiert“, d.h. verstärkt auf den sozialen Kontext und dessen Bedeutung bezogen.

Neben den Hinweisen, die für die Gültigkeit der Theorie der sozialen Identität sprechen, gibt es auch einige Problemstellungen, die bisher nicht mit den theoretischen Annahmen vereinbar sind. Hierzu zählen zum einen Befunde zum Zusammenhang zwischen Gruppenidentifikation und Eigengruppenfavorisierung: Es wird angenommen, dass ein direkter Zusammenhang zwischen diesen beiden Konstrukten eine notwendige Voraussetzung für die Gültigkeit der Theorie darstellt, dem gegenüber aber zahlreiche Befunde stehen, welche einen solchen Zusammenhang nicht oder nur in unbedeutendem Maße feststellen (Brown, Hinkle, Ely, Fox-Cardamone, Maras & Taylor, 1992), was als Argument gegen die Gültigkeit der Theorie angeführt wird. Turner (1999) diskutiert diese Problematik umfassend und argumentiert u.a., dass eine Eigengruppenfavorisierung lediglich *eine* Strategie zur Herstellung positiver Distinktheit darstellt und somit die Theorie der sozialen Identität den *ingroup bias* keinen entsprechenden, zwingenden Zusammenhang zwischen Eigengruppenfavorisierung und Gruppenidentifikation herstellt, sondern lediglich eine möglicherweise mediierende Wirkung unterstellt.

Weiterhin wird unter Bezugnahme auf die „*self-esteem*-Hypothese“ diskutiert, dass die Diskrimination von Außengruppen den eigenen Selbstwert steigert und das Bedürfnis nach einem positiven Selbstwert im Falle einer Selbstwertbedrohung dazu führt, dass diskriminierendes Verhalten zwischen Gruppen wahrscheinlicher wird (Abrams & Hogg, 1990). Demzufolge sollte ein Zusammenhang zwischen der Eigengruppenfavorisierung und individuellem Selbstwert bestehen, resp. sollte ein geringer, individueller Selbstwert mit einer verstärkten Eigengruppenfavorisierung einhergehen. Dies sichert im Sinne der Theorie ab, dass man die eigene Gruppe und damit auch sich selbst als Mitglied dieser Gruppe positiv wahrnehmen kann. Der

erste Zusammenhang (diskriminierendes Intergruppenverhalten erhöht individuellen Selbstwert) konnte empirisch bisher vergleichsweise gut belegt werden, wohingegen der zweite Zusammenhang (deprivierter Selbstwert führt zu verstärkter Intergruppendifferenzierung) empirisch bisher wenig Unterstützung erhalten hat – insbesondere durch die verstärkte Messung von Selbstwert als globalem Faktor im Sinne eines *trait*, wohingegen eine situationspezifische Erfassung als *state* den Annahmen der Theorie der sozialen Identität eher entsprechen würde. Ferner wird subsumiert, dass intergruppal Diskriminierung zwar den Selbstwert erhöht, aber nicht durch ein Bedürfnis der Selbstwerterhöhung motiviert zu sein scheint, was der Theorie widerspricht (Rubin & Hewstone, 1998). Dies führt ferner bis zu der Annahme, dass der direkte, motivationale Einfluss des Selbstwertes auf das Zustandekommen von (diskriminierendem) Intergruppenverhalten, wie er in der ursprünglichen Theorie postuliert wird, angezweifelt werden kann und eher einen Nebeneffekt darstellt (Hogg & Mullin, 1999).

Das Phänomen der positiv-negativ Asymmetrie sozialer Diskriminierung (Mummendey et al., 1992) beschreibt die Tatsache, dass im *minimal group*-Paradigma dann, wenn *In- und Outgroup* anhand positiver Dimensionen (wie z.B. Intelligenz) bewertet werden oder positiv besetzte Ressourcen (wie z.B. Geld) zugewiesen werden eine Eigengruppenfavorisierung, resp. Außengruppendiskriminierung stattfindet. Beim Einsatz negativ konnotierter Dimensionen (z.B. Dummheit, Langeweile) oder negativer Behandlungen (z.B. Dauer der Applikation eines unangenehmen Tones) bleibt hingegen die Diskriminierung der Außengruppe aus (vgl. auch Mummendey, Otten, Berger & Kessler, 2000). Somit scheint ein und derselbe Intergruppenkontext, welcher die hinreichende Bedingung für soziale Diskriminierung im positiven Bereich darstellt, nicht ebenso im negativen Bereich zu funktionieren, was der Theorie der sozialen Identität widerspricht, da diese keine Unterscheidungen hinsichtlich der Valenz von Intergruppenverhalten macht, diese aber empirisch nachgewiesen sind (Buhl, 1999). Kritisch diskutiert wird dennoch, ob durch die Anlage des Experiments von Mummendey et al. (1992) eine kognitive Restrukturierung des experimentellen Settings und damit auch einen Wandel der Gruppenkonstellation herbeigeführt hat, in dem Sinne, dass nun die Probanden/-innen des Experiments eine gesamte *Ingroup* bilden und nunmehr die Testleiter/-innen die *Outgroup* darstellen, was die Reduzierung der Außengruppendiskriminierung (da die Testteilnehmenden nunmehr eine einzelne Gruppe bilden) im Sinne der Theorie der sozialen Identität erklären könnte (Gärtner, Dovidio, Anastasio, Bachman & Rust, 1993).

Zuletzt weisen Befunde darauf hin, dass entgegen der Annahmen der *Social Identity Theory*, die von einer maximalen Differenzierung zwischen Gruppen ausgeht, Effekte einer wahrgenommenen Ähnlichkeit zwischen zwei Gruppen auftreten können. Nach der Theorie wäre zu erwarten, dass sich besonders Gruppen, die sich als ähnlich wahrnehmen, bemühen, sich maximal voneinander zu differenzieren (Turner, 1978). Empirisch zeigt sich allerdings, dass Ähnlich-

keiten zwischen Gruppen nicht im Sinne der Theorie zu differenzierendem Intergruppenverhalten führen, sondern eine höhere Intergruppenattraktivität, eine höhere Kontaktbereitschaft mit der *Outgroup* sowie ein verringerter *ingroup bias* die Folgen des Intergruppenvergleichs darstellen. Auf für die Gruppendifinition hoch relevanten Merkmalen zeigte sich zwar eine Eigengruppenbevorzugung, wohingegen auf hierfür weniger relevanten Merkmalen eine Außengruppenbevorzugung zu finden war (Grant, 1993; Roccas & Schwartz, 1993). Eine Bevorzugung der Außengruppe ist jedoch mit den Grundannahmen der Theorie der sozialen Identität auch auf vergleichsweise peripheren Vergleichsdimensionen nicht vereinbar und wird daher als Argument gegen die umfassende Gültigkeit der Theorie angeführt. Brewer (1991) zeigt diesbezüglich, dass Depersonalisation und Gruppengröße als Determinanten sozialer Identität miteinander interagieren und schlägt ein Modell „optimaler Distinktheit“ vor, welches einen Einbezug von Assimilation und Differenzierung als Bedürfnisse von Intergruppenverhalten postuliert. Demnach ist für die Konstitution sozialer Identität nicht ausschließlich eine maximale Differenzierung von der Fremdgruppe maßgeblich, wie von der *Social Identity Theory* vorgeschlagen, sondern auch die Wahrnehmung von Gemeinsamkeiten innerhalb der Eigengruppe von zentraler Bedeutung (sog. „*group loyalty*“). Nur auf diese Weise können beide Grundbedürfnisse nach Unterschiedlichkeit/Individualität (durch den Vergleich mit Fremdgruppenmitgliedern) und Ähnlichkeit/Zugehörigkeit (durch den Vergleich mit Eigengruppenmitgliedern) befriedigt werden (Brewer & Pickett, 2002). Die ausschließliche Fokussierung auf eine maximale Differenzierung zwischen Gruppen nach der Theorie der sozialen Identität reicht demnach nicht für die Erklärung von Intergruppenprozessen aus, vielmehr ist neben maximaler Differenzierung eine weitere Facette – das Bedürfnis nach Ähnlichkeit – konstitutiv für die soziale Identität, was bisher in der Theorie der sozialen Identität nicht angemessen berücksichtigt wird.

Zusammenfassend lässt sich festhalten, dass die Theorie der sozialen Identität die Forschung zu Intergruppenprozessen auf vielfache Weise befruchtet und entscheidend geprägt hat. Besonders hervorzuheben bleibt, dass die Theorie die Analyse von sozialem Verhalten wieder „resozialisiert“, d.h. sich von einer Übertragung von Theorien und Konzepten zu interpersonalem Verhalten auf Gruppen anwendet, was Tajfel (1972) bereits forderte.

Wenngleich eine umfassende, empirische Absicherung aller Annahmen der Theorie bis zum jetzigen Zeitpunkt nicht möglich war, so können jedoch die Kernannahmen als empirisch abgesichert bezeichnet werden. Trotz einiger empirischer Befunde, welche nicht umfassend mit der Theorie in Einklang gebracht werden können, legen diese nicht die Ungültigkeit der gesamten Theorie nahe, sondern vielmehr die Notwendigkeit, an einigen Stellen die theoretischen Annahmen differenzierter zu betrachten.

Aus diesen Gründen kann die Theorie der sozialen Identität im Zusammenhang mit der Fragestellung der vorliegenden Arbeit somit als wertvolle Erklärungsgrundlage für Intergrup-

penverhalten von Jungen und Mädchen im schulischen Kontext herangezogen werden. Im folgenden Abschnitt soll dargelegt werden, wie die theoretischen Annahmen der *Social Identity Theory* auf das Anwendungsfeld von Jungen in der Schule übertragen werden können und welche Verhaltenskonsequenzen der Theorie folgend im Anwendungskontext erwartet werden können.

4.5. Anwendung auf die Problemstellung

Die einführend dargestellte, anhaltende öffentliche Debatte über das Versagen von Jungen in der Schule führt bei Jungen im Sinne der Theorie der sozialen Identität zu einem ungünstigeren Vergleichsergebnis mit der relevanten *Outgroup*, den noch zur Schule gehenden Mädchen. Damit entsteht für Jungen eine bedrohte soziale Identität sowie damit einhergehend das Bedürfnis, die positive Distinktheit zur Außengruppe wiederherzustellen. Bei einer starken Vergleichbarkeit zwischen den Gruppen ist der Theorie der sozialen Identität folgend wahrscheinlich, dass sozialer Wettbewerb zunimmt, je stärker die Vergleichbarkeit zwischen den Gruppen zu bewerten ist. In einer Gesellschaft, in denen Geschlechtergleichheit rechtlich abgesichert und zusätzlich öffentlich betont werden (Frevert, 1999), kann angenommen werden, dass männliche und weibliche Identitäten auf einer großen Anzahl gemeinsamer Wertedefinitionen basieren. Dies wiederum hat aber zur Folge, dass Geschlecht als unterscheidende Variable salient wird und für die Herstellung positiver Distinktheit eine zentrale Rolle einnimmt. Somit wird es notwendig, dass beide Gruppen auf Basis der Geschlechtszugehörigkeit danach streben, sich als positiv distinkt wahrzunehmen – es liegt faktisch die im Rahmen der Theorie postulierte negative Interdependenz zwischen den Geschlechtern vor. Auf den Schulkontext appliziert entsteht somit eine nachteilige Lage für Jungen, da die schulische Überlegenheit dieser Gruppe zwar in der Vergangenheit Geltung zu haben schien (Statistisches Bundesamt, 2014b), nun aber einem Wandel unterliegt und durch ein positiveres Abschneiden von Mädchen ein bedrohlicher sozialer Vergleich entstehen kann. Die hieraus resultierende, ungesicherte soziale Identität muss daher nun durch die o.g. Strategien wieder abgesichert werden.

Es ist somit denkbar, dass Jungen getriggert durch die Wahrnehmung eines Wandels der bisher geltenden Intergruppenbeziehung (und besonders bei Bewertung dieses Wandels als instabil und illegitim) vermehrt Strategien einsetzen, um die positive Distinktheit weiterhin aufrechtzuerhalten bzw. wiederherzustellen. Ein Wechsel der Gruppe scheidet aus, da das definierende Merkmal Geschlecht als stabil im Sinne der Theorie der sozialen Identität wahrgenommen wird und daher keine Möglichkeit des Gruppenwechsels zur Stabilisierung der bedrohten sozialen Identität antizipiert werden kann. Die Theorie bietet als Strategien des direkten sozialen Wettbewerbs an, einerseits vermehrt Verhaltensweisen zu zeigen, welche die Überlegenheit der *Ingroup* besonders betonen (denkbar wären u.a. die Inszenierung größerer körperlicher Stärke

(Stauber, 2012), oder die Betonung noch immer bessere Karrierechancen von Männern (Holst & Wiemer, 2010; Leuze & Strauß, 2009). Andererseits besteht laut Theorie eine zweite Möglichkeit, die Abwertung der *Outgroup*: Da Geschlecht, wie oben beschrieben, eine zentrale Vergleichsdimension des beschriebenen Intergruppenkonflikts darstellt, kann geschlussfolgert werden, dass im Speziellen also Weiblichkeit abgewertet wird (z.B. durch sexistische Abwertungen (Frazier & Sadker, 1973) oder die Unterrepräsentanz von Frauen in gesellschaftlich und wirtschaftlich bedeutsamen Berufspositionen (Holst & Wiemer, 2010; Heisig, 2011) herauszustellen, um die positive Distinktheit der Eigengruppe der Jungen wiederherzustellen.

Im Rahmen der „sozialen Kreativität“ können Jungen neue Vergleichsdimensionen erfinden, auf welchen Sie wiederum besser abschneiden als Mädchen. So wäre denkbar, dass Jungen statt schulischem Erfolg verstärkt auf dessen Umsetzung in eine berufliche Karriere fokussieren oder den schulischen Kontext nicht mehr als Vergleichsdimension heranziehen, sondern vielmehr außerschulische Aktivitäten (z.B. Vereinsaktivität: Keil, 2013; Sportlichkeit: Wissenschaftliches Institut der Ärzte in Deutschland, 2000) als Vergleichsmaßstab anwenden, bei denen die *Ingroup* positiver abschneidet. Als zweite sozial kreative Strategie bietet sich die Änderung der Vergleichsgruppe an. Jungen im Kontext Schule werden sich dabei der Theorie folgend am ehesten an den vergleichbarsten verfügbaren Mitgliedern der *Outgroup* orientieren, was im Klassenverband für einen Vergleich mit Mädchen spricht. Fällt dieser allerdings negativ aus bzw. scheint die bisher wahrgenommene Überlegenheit mit dieser spezifischen Vergleichsgruppe bedroht, kann eine alternative Vergleichsgruppe herangezogen werden, so z.B. Mädchen (oder auch Jungen) einer schwächeren Klasse oder einer niedrigeren Schulform. Abschließend ist die Möglichkeit denkbar, dass eine Umbewertung von geringem schulischem Erfolg als positiv stattfindet, also dass Jungen *school loafing* positiver bewerten und es als wünschenswert erachten, den schulischen Anforderungen nicht (vollumfänglich) gerecht zu werden. Über das damit verbundene, negativere Abschneiden der leistungsstarken Mädchengruppe hinsichtlich dieser umbewerteten Vergleichsdimension lässt sich für Jungen ebenso die bedrohte soziale Identität stabilisieren. Zu beachten ist, dass die Wiederherstellung der positiven Distinktheit hierbei stark vom Realitätsbezug des Vergleiches abhängt und ein Vergleich mit zu fernen Vergleichsdimensionen oder einer enorm statusniedrigeren Gruppe trotz eines positiven Vergleichsergebnisses nicht zu einer Stabilisierung der bedrohten sozialen Identität führen kann (bspw. Vergleiche von Jungen auf dem Gymnasium mit Mädchen einer Förderschule). Es ist bei dieser Strategie also zentral, die Vergleichsgruppe zwar so stark wie nötig, aber so wenig wie möglich zu verändern, um dieser Problematik zu entgehen.

Es konnte somit verdeutlicht werden, dass die Theorie der sozialen Identität angewendet auf die Darstellung von Jungen als Bildungsverlierer schlüssig erklären kann, warum Jungen, die mit der medialen und öffentlichen Darstellung ihres schulischen Scheiterns konfrontiert werden, eine Bedrohung der eigenen sozialen Identität wahrnehmen sollten. Des Weiteren schlägt die

Theorie mögliche Strategien vor, wie die bedrohte soziale Identität wiederhergestellt und positive Distinktheit zur *Outgroup* erreicht werden kann. Somit liefert die Theorie der sozialen Identität eine wichtige Grundlage für das Verstehen von Intergruppenprozessen zwischen Jungen und Mädchen im Kontext Schule, welche insbesondere auch durch den stattfindenden Paradigmenwechsel, d.h. einer erstarkenden Diskussion um Bildungsnachteile von Jungen (Weaver-Hightower, 2003), eine empirische Prüfung zur Anwendbarkeit der Theorie bei Jungen im schulischen Kontext erkennen lassen.

II. EMPIRISCHER TEIL

Alle in diesem Kapitel präsentierten, empirischen Untersuchungen von Schüler/-innen wurden durch die Berliner Senatsverwaltung für Bildung, Wissenschaft und Forschung genehmigt.

5. Studie 1: Wie werden Jungen wahrgenommen? Zur Existenz des Stereotyps des bildungsverlierenden Jungen bei Schüler/-innen

Veröffentlicht als Studie 1 in:

Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!? How the public discourse on failing boys impacts performance outcomes and motivational goal orientation in German students. *Social Psychology, 45*, 112-126.

Abstract. We investigated effects of the media's portrayal of boys as "scholastic failures" on secondary school students. Experimental exposure to the negative portrayal induced both stereotype threat and reactance. Stereotype threat appeared through boys' underperformance in reading but not in mathematics. Stereotype reactance was evident as boys displayed stronger learning goals towards mathematics but not reading. Apparently, boys were motivated to disconfirm their group's negative depiction, however, while they could successfully apply compensatory strategies when describing their learning goals, this motivation did not enable them to perform better. The media portrayal thus contributes to the maintenance of gender-stereotypes, by impairing boys' performance in female connoted domains and by prompting them to align their learning goals to the gender-connotation of the domain.

Key words: stereotype threat, stereotype reactance, gender stereotypes, gender differences, "failing boys"

5.1. Überblick und Fragestellung

Wie bereits eingangs dargelegt wurde, bildet die mediale Darstellung von Jungen als neue Verlierer des Bildungssystems nur einen Ausschnitt aus der empirischen Realität ab (vgl. Kapitel 1). Es finden sich in den zwei männlich konnotierten Kompetenzbereichen der PISA-Studie zwar nur geringe bis keine (nach den Geschlechterstereotypen erwartbare) Leistungsvorteile, aber auch *keine Nachteile* für Jungen. Die einzige Kompetenzeinbuße zuungunsten von Jungen in der Lesekompetenz ist zwar nicht minder problematisch für Jungen, da Lesen eine der wichtigsten Vorläuferkompetenzen für andere akademische Bereiche ist, aber auch unter Berücksichtigung dieser Tatsache ist die mediale Darstellung von Jungen als Schulversager im Allgemeinen deutlich verzerrt. Weitere Forschungsergebnisse deuten ergänzend darauf hin, dass wenn überhaupt nachteilige Effekte für Jungen in der Schule nachgewiesen werden, die Effektstärken der Befunde eher gering ausfallen (insbesondere bei Kontrolle der Schulform).

Es kann somit festgehalten werden, dass entgegen der verbreiteten, medialen Darstellung von Jungen als Bildungsverlierer insgesamt dieses Bild bei Weitem nicht so generalisiert vorzufinden ist, wie entsprechende Veröffentlichungen suggerieren. Während Daten der großen Schulleistungsstudien Nachteile von Jungen im Lesen finden, deuten andere Ergebnisse darauf hin, dass es bei Kontrolle der Schulform auch hier keine nennenswerten Nachteile für Jungen gibt. Weitere, mehr am Curriculum und weniger kompetenzorientierte Arbeiten ergänzen diese Befunde und bestätigen ebenfalls, dass Jungen gar nicht oder zumindest nicht generell als Bildungsverlierer dargestellt werden können.

Trotzdem sind Medien als Instrumente von Unterhaltung und Meinungsbildung auch bei Jugendlichen allgegenwärtig und stellen einen bedeutenden Anteil am Alltagsleben von ihnen dar. Insbesondere dem Internet kommt hier eine besondere Bedeutung zu (Klingler, 2008). Hinzu kommt, dass jüngere Nutzer von Medien erst entsprechende Kompetenzen im Umgang mit Medien erlernen (Medienkompetenz entwickeln) müssen, ehe sie in der Lage sind, entsprechend auf mediale Darstellungen zu reagieren, diese kritisch und angemessen zu bewerten und richtige Schlussfolgerungen aus ihnen abzuleiten (vgl. Treumann et al., 2007). Die durch die Medien transportierten und bei Printmedien meist im Internet vollständig dokumentierten Inhalte gelangen so möglicherweise auch zu jugendlichen Medienkonsumentinnen und -konsumenten und können dann maßgeblich zur Entstehung und/oder Aufrechterhaltung von Stereotypen beitragen (Appel, 2008).

Da aufgrund bisheriger Forschungsergebnisse nicht beantwortet werden kann, inwieweit die mediale Darstellung von Jungen als Bildungsverlierer auch in der Wahrnehmung von Jugendlichen angekommen ist, soll es Ziel dieser Studie sein, zu überprüfen inwieweit Schülerinnen und Schüler durch den medialen Diskurs von Jungen als Bildungsverlierern eine entsprechende gesellschaftliche Meinung von generell schulversagenden Jungen antizipieren, oder sie eher ge-

schlechterstereotype Annahmen über die schulische Leistung von Jungen reproduzieren. Hierzu wurde untersucht, inwieweit Schülerinnen und Schüler davon ausgehen, dass die gesellschaftliche Meinung durch ein Bild von Jungen als schlechtere Schüler geprägt ist. Die Jugendlichen wurden einerseits gebeten, freie Äußerungen zur Wahrnehmung (ohne *framing* auf einen bestimmten Bereich) zu produzieren, als auch anschließend das Verhalten von Jungen im Allgemeinen und spezifisch in der Schule auf semantischen Differentialen zu bewerten. Als Ergebnis dieser Studie liegen somit qualitative, frei produzierte und quantitative Daten vor, welche sich wechselseitig ergänzen und einen umfassenden Einblick in die Meinung von Jugendlichen über ihre gesellschaftliche Wahrnehmung im Allgemeinen und in der Schule ermöglichen. Neben der Besonderheit eines qualitativen und eines quantitativen Zugangs fokussiert die vorliegende Studie nicht explizit nur auf die antizipierte Wahrnehmung von Jungen über ihre eigene *Ingroup*, sondern zieht auch ergänzend den Blickwinkel von Mädchen zusätzlich heran.

Es kann diskutiert werden, ob im Rahmen der folgenden Studie *Metastereotype*, also die die Annahmen, welche *Stereotype* einer *Outgroup* von der Eigengruppe (*Ingroup*) antizipiert werden (z.B.: „Was denkt die Gesellschaft über Schüler/-innen?“), Untersuchungsgegenstand sind oder aber aufgrund des großen, auf die gesamte Gesellschaft ausgedehnten Referenzrahmens („die Leute“) nicht mehr von einer klar abgrenzbaren *Outgroup* gesprochen werden kann und somit generelle stereotype Wahrnehmungen zum Untersuchungsgegenstand werden. Da unter Bezugnahme auf die aktuelle Forschungsliteratur keine eindeutige Positionierung möglich ist (z.B.: Gordijn, 2002 und Gomez, 2002 bieten eine sehr weite Definition an [„*the others*“; Gordijn, 2002, S. 283; Gomez, 2002, S. 253], wohingegen Vorauer, Main & O’Connell (1998) sowie Vorauer, Hunter, Main & Roy (2000) *Metastereotype* auf *eine* umschriebene Gruppe abgrenzen [„*particular out-group*“; Vorauer et al., 1998, S. 917; Vorauer et al., 2000, S. 690]) werden im Folgenden beide Begrifflichkeiten gemeinsam verwendet.

Diese Studie liefert ferner eine wichtige Grundlage für weitere Untersuchungen im Rahmen dieser Arbeit, für welche die Erkenntnis zentral ist, ob es bereits ein Stereotyp von „Jungen als Schulversager“ gibt. Im Folgenden sollen nun das Vorgehen und die Ergebnisse dieser Untersuchung zur antizipierten, gesellschaftlichen Wahrnehmung von Jungen (und Mädchen) erläutert werden.

Untersuchungsgegenstand der vorliegenden Studie ist somit, ob Schülerinnen und Schüler bereits (Meta-)Stereotype über die Gruppe der Jungen im Allgemeinen und insbesondere hinsichtlich des schulischen Erfolgs von Jungen ausgebildet haben. Um dies festzustellen wurden Jungen und Mädchen über antizipierte, gesellschaftliche stereotypische Vorstellungen über Jungen im Allgemeinen und in der Schule (was „die Leute“ denken) sowohl mit einem offenen Antwortformat als auch auf verschiedenen semantischen Differentialen befragt.

Folgende Untersuchungshypothesen werden formuliert und sollen im Rahmen dieser Studie geprüft werden:

- H1* Jungen antizipieren ein positiveres, allgemeines gesellschaftliches Bild ihrer Eigengruppe und produzieren mehr allgemein positive Äußerungen, als Mädchen dies für die Fremdgruppe der Jungen tun (*ingroup bias*).
- H2* Mädchen antizipieren ein positiveres, allgemeines gesellschaftliches Bild ihrer Eigengruppe und produzieren mehr allgemein positive Äußerungen, als Jungen dies für die Fremdgruppe der Mädchen tun (*ingroup bias*).
- H3* Jungen und Mädchen produzieren mehr negative Äußerungen über Jungen in der Schule, als über Mädchen.

5.2. Methode

Teilnehmende. Für die Untersuchung konnten 206 Schüler/-innen, davon 117 Mädchen und 89 Jungen der 9. Klassenstufe zweier Berliner Gesamtschulen gewonnen werden. Das Alter betrug zum Zeitpunkt der Messung ca. 14.5 Jahre ($M = 14.67$, $SD = 0.67$, *range* 14-17 Jahre). Um den Besonderheiten des Erhebungsorts Berlin insbesondere im Hinblick auf die relativ hohe Streuung beim sozioökonomischen Status und einen vergleichsweise hohen Anteil an Familien mit Migrationshintergrund gerecht zu werden, entstammen die beiden Partnerschulen für diese Untersuchung Bezirken, welche in Sozioökonomie und Migrationsanteil dem bundesdeutschen Durchschnitt nahe kommen (Spandau, Steglitz). Zur Überprüfung der Verteilung beider Kennwerte Migrationshintergrund und sozioökonomischer Status wurden zentrale Variablen erfasst, wie das Geburtsland der Schülerinnen und Schüler sowie deren Eltern (Deutschland vs. anderes Land) und die Sprache, welche die Schülerinnen und Schüler am meisten mit verschiedenen Bezugspersonen sprechen (Eltern, Geschwister, Verwandte wie z.B. Großeltern oder Onkel/Tante sowie Freunde), um trotz (v.a. in den folgenden Studien) stattgefundenener randomisierter Zuweisung auf verschiedenen Bedingungen die Gleichverteilung zentralster Hintergrundvariablen überprüfen zu können. Der sozioökonomische Hintergrund wurde mit der Frage nach (bildungsrelevanten) heimischen Besitztümern (Auswahl aus „home possessions“, Fragennummern entsprechend des PISA-Skalenhandbuchs 2006; Frey et al., 2009: ST20Q01, Q02, Q03, Q10, Q11, Q12) operationalisiert³.

³ Es war angedacht, sowohl Migrationshintergrund als auch den sozioökonomischen Status exakter und differenzierter zu erheben, u.a. mit Fragen nach dem genauen Geburtsland, der Ankunft in Deutschland sowie nach dem Bildungsabschluss der Eltern, der Anzahl von Büchern („Bücherfrage“) und weiteren Kulturgütern im Haus sowie dem Beruf / Erwerbsstatus der Eltern. Derartige Angaben durch die Schülerinnen und Schüler einzuholen wurde durch die genehmigende Behörde untersagt. Auf die Herausgabe eines separaten Elternfragebogens wurde verzichtet, da unter Berücksichtigung der üblichen Rück-

Die Angaben zum Geburtsland und zur gesprochenen Sprache mit verschiedenen Bezugsgruppen sind in der folgenden Tabelle zusammengefasst:

Tabelle 3: Direkte Migrationserfahrung und gesprochene Sprache/n getrennt nach Geschlecht. $N = 206$, davon 5 fehlende Angaben zum Geburtsland und 2 fehlende Angaben zur Sprache.

	Geburtsland		Sprache mit Eltern	
	Deutschland	and. Land	Deutsch	and. Sprache ⁴
Mädchen	$n = 106$	$n = 11$	$n = 48$	$n = 68$
Jungen	$n = 72$	$n = 12$	$n = 36$	$n = 52$
Gesamt	$n = 178$	$n = 23$	$n = 84$	$n = 120$

Im Hinblick auf die Quote für direkt zugewanderte Kinder entspricht diese in etwa dem Bundesdurchschnitt, Kinder mit Migrationshintergrund in der zweiten Generation (operationalisiert über die Sprache mit den Eltern: ausschließlich andere Sprache) sind trotz sorgfältiger Auswahl der Schulen gemessen am bundesdeutschen Durchschnitt leicht überrepräsentiert (vgl. Stanat, Rauch & Segeritz, 2010). Für die vorliegende Untersuchung bedeutsam ist, dass sich beide Geschlechter weder hinsichtlich des Geburtslandes, noch nach der gesprochenen Sprache bedeutsam voneinander unterscheiden, $\chi^2 < 1.15$, $p > .28$.

Wird die Sprache als Indikator für Migrationserfahrung betrachtet, zeigt sich ein deutlich erhöhter Anteil von insgesamt 67.80% Personen, die eine nicht-deutsche Sprache mit einer Bezugsgruppe (siehe Fußnote 3) sprechen. Möglicherweise liegt bei diesem Personenkreis eine Zuwanderung in einer früheren Generation (indirekte Migrationserfahrung) vor. Da hierzu keine belastbaren Angaben vorliegen, sind diese Angaben als Indikator für Zuwanderungshintergrund allerdings mit Vorsicht zu interpretieren.

Im Mittel geben die befragten Schüler/-innen an, dass ihnen 5 von 7 erfragten heimischen Besitztümern zur Verfügung stehen ($M = 5.32$, $SD = 1.33$; keine Unterschiede nach Geschlecht mit: $M_{Jungen} = 5.33$, $SD = 1.47$ und $M_{Mädchen} = 5.31$, $SD = 1.21$; $F < 1$, n.s.). Gemessen an der Stichprobe von PISA 2009 für Deutschland ergibt sich hieraus ein vergleichbar großer sozioökonomischer Status der untersuchten Gruppe (OECD, 2011).

Prozedur und Material. Die Erhebung fand in Verbindung mit weiteren Erhebungen (Studien 4 und 5) in zwei aufeinanderfolgenden, normalen Schulstunden im gesamten Klassenverbund statt. Alle anwesenden Schüler/-innen wurden nach der Begrüßung über den Ablauf der

laufquoten die Anzahl an benötigten Versuchspersonen enorm gesteigert werden müsste. Ferner ergibt sich die Problematik, dass Eltern- und Schüler/-innenfragebogen eindeutig zuordenbar sein müssen, was das die zugesicherte Anonymität der Befragung potentiell gefährden kann.

⁴ Zusätzlich erfragt wurde die Kommunikation mit den Personengruppen: Geschwister, Verwandte, Freunde. Wird als Indikator die Kommunikation mit mind. einer dieser Personengruppen *oder* den Eltern gebildet ergibt sich ein etwas niedrigerer Anteil der Verwendung der deutschen Sprache (Mädchen: $n = 39$ deutsch, $n = 78$ andere Sprache; Jungen: $n = 27$ deutsch, $n = 61$ andere Sprache; 1 fehlende Angabe).

Untersuchung informiert und über deren Rahmenbedingungen wie Freiwilligkeit der Teilnahme und die vollständig anonyme Datenerhebung aufgeklärt. Bereits zu diesem Zeitpunkt wurden – wie in allen anderen im Rahmen dieser Arbeit vorgestellten Studien – klare Regeln für den Untersuchungsablauf etabliert, z.B. dass ähnlich einer Klassenarbeit Gespräche zwischen den Schüler/-innen und gegenseitiges Beobachten während des Ausfüllens der Testhefte unerwünscht seien. Alle anwesenden Schüler/-innen willigten daraufhin in eine Teilnahme an der Untersuchung ein. In ausbalancierter Abfolge wurden die Teilnehmer/-innen dann gebeten, zu antizipieren, was das gesellschaftliche Meinungsbild über a) Jungen und b) Mädchen beinhaltet und dies frei niederzuschreiben (siehe Anhang). Die genaue Instruktion lautete wie folgt:

*„Es geht nun darum, was du glaubst, was die Leute über Jungen [alternativ: Mädchen] denken. Schreibe bitte auf, was **die Leute** über Jungen [alternativ: Mädchen] denken, auch wenn du darüber selbst anders denkst! Schreibe so viele Wörter oder Sätze auf, wie dir einfallen. Es gibt hierbei keine falschen oder richtigen Antworten“*

Um die schriftliche Instruktion zu ergänzen und leichter verständlich zu machen, wurde durch einen standardisierten Untersuchungsleitfaden mündlich Folgendes ergänzt:

„Es ist wichtig, dass ihr hier nicht eure eigene Meinung aufschreibt, sondern dass, was die Leute denken. Also wenn ihr zum Beispiel auf der Straße einfach jemanden ansprecht und fragt „Was denken Sie, wie sind Jungs?“ – welche Antworten würdet ihr dann erwarten? Genau das, was ihr dann als Antwort erwarten würdet, sollt ihr aufschreiben.“

Daraufhin wurde den Teilnehmer/-innen 10 Minuten Zeit gewährt, um so viele Antworten wie möglich zu generieren. Wurde bei einzelnen Teilnehmenden bereits vor Ablauf dieser Zeit die Bearbeitung beendet, wurden diese dazu angehalten möglichst noch weitere Antworten zu ergänzen. Nach Ablauf der Zeit wurde die Bearbeitung der Aufgabe unterbrochen. Die Messung stereotyper Annahmen in diesem Teil der Untersuchung erfolgte auf eine indirekte Art und Weise, da nicht auf die persönliche Meinung, sondern eine übergeordnete Metaebene fokussiert wurde.

Im Anschluss an die freie Äußerung wurde den Teilnehmer/-innen ein Fragebogen mit 28 semantischen Differentialen (5stufig, Auszug aus Knigge & Hannover, 2011) präsentiert, auf welchem angekreuzt werden sollte, wie die typische Meinung der Bevölkerung über Jungen a) im Allgemeinen und auf einem separaten Bogen b) in der Schule vermeintlich ausfallen würde (siehe Anhang). Der Fragebogen bestand aus leistungsorientierten Eigenschaften (z.B. leistungsfähig vs. leistungsunfähig) und Adjektiven zum Sozialverhalten (z.B. vorlaut vs. zurückhaltend).

Zur Bearbeitung dieses Testteils standen ebenfalls 10 Minuten zur Verfügung.

5.3. Ergebnisse

Zwei unabhängige *Rater* extrahierten aus den von Schülerinnen und Schülern produzierten offenen Angaben zunächst alle Verhaltensbeschreibungen (wie z.B. „Jungs sind laut und stören den Unterricht“) und alle Persönlichkeitseigenschaften im Sinne von *traits* (z.B. „Jungen sind intelligent“) als Bestandteile von (Meta-)Stereotypen über Jungen und Mädchen. Eine domänen-spezifische Auswertung getrennt nach Schulfächern (z.B. Deutsch und Mathematik), wie ursprünglich geplant, war leider nicht möglich, da weniger als 1% der freien Äußerungen auf ein konkretes Schulfach bezogen wurden. Sämtliche synonyme Äußerungen einer Testperson (z.B. ungebildet, dumm und unwissend) wurden als lediglich eine Äußerung erfasst.

Unter dieser Maßgabe resultierte eine Gesamtzahl von $N = 961$ Aussagen. Jede der offenen Angaben wurde in einem weiteren Schritt von beiden *Ratern* zunächst als positiv oder negativ valent eingestuft. Darauf aufbauend wurden die bereits kategorisierten Äußerungen dahingehend überprüft und kodiert, ob ein klarer Lern- oder Schulbezug erkennbar ist (z.B. dumm, strebsam usw.) oder nicht (z.B. einfühlsam, aufgeschlossen) in den Aussagen zu erkennen war. Die Interrater-Übereinstimmung lag mit Cohen's $\kappa = 0.86$, und einer Übereinstimmung von 93.4% in einem guten bis sehr guten Bereich (Cohen, 1960).

Von allen kodierten Angaben entfallen 426 negative Äußerungen und 87 positive auf Jungen, wohingegen bei Mädchen mit 211 negativen Angaben im Vergleich zu 237 positiven ein eher ausgeglichenes Meinungsbild vorherrscht. Nichtparametrische Analysen mit dem Chi-Quadrat-Test deuten auf einen signifikanten Unterschied zwischen den auf Jungen vs. auf Mädchen bezogenen Angaben hin, $\chi^2(1, N = 961) = 138.24, p < .001$. Im Hinblick auf die Angaben mit einem eindeutigen Schulbezug dominiert mit 165 negativen und 21 positiven Kommentaren eine deutlich negative Sichtweise im Bezug auf Jungen in der Schule. Betrachtet man hingegen die Angaben zu Mädchen in der Schule kehrt sich dieses Bild um: Hier werden 14 negative und 136 positive Eigenschaften durch die Schülerinnen und Schüler genannt. Auch diese Differenzen sind mit $\chi^2(1, N = 336) = 185.91, p < .001$ statistisch bedeutsam. Einen Überblick über die Verteilung aller Antworten von Jungen und Mädchen gibt die folgende Abbildung:

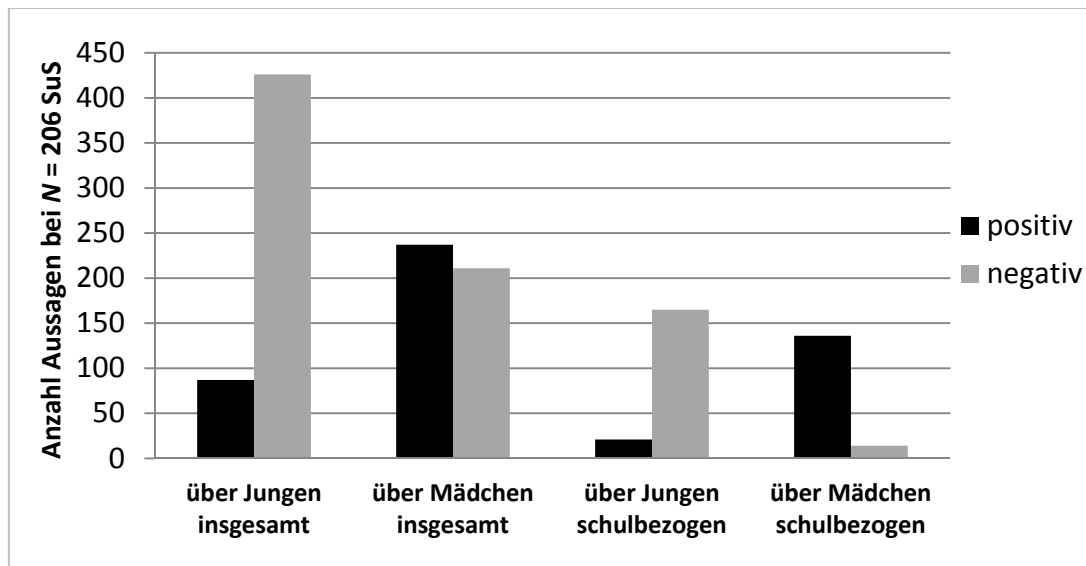


Abbildung 3: Positive und negative Aussagen allgemein sowie mit konkretem Schulbezug von $N = 206$ Schüler/-innen, nach Kontext (insgesamt vs. Schule) und Valenz (positiv vs. negativ).

Analysen zum *ingroup bias* zeigen bei durch Jungen gemachten Angaben über die Eigengruppe 161 generell negative und 33 positive Aussagen – hier überwiegen somit trotz der Beurteilung der Eigengruppe negative Eigenschaften, welche Jungen über die eigene Gruppe als gesellschaftliches Bild antizipieren. Mädchen charakterisieren die Gruppe der Jungen mit 265 negativen im Vergleich zu 54 positiven Eigenschaften mit einer vergleichbar hohen Quote an negativen Äußerungen (Jungen über Jungen: 82.99% negative Äußerungen; Mädchen über Jungen: 83.07% negative Äußerungen). Bei den Angaben von Mädchen über die Eigengruppe finden sich 143 negative und 133 positive Aussagen, Jungen produzieren diesbezüglich 68 negative und 104 positive Äußerungen, $\chi^2(1, N = 448) = 6.41, p = .01$. Die Quoten der negativen Zuschreibungen sind hierbei mit 50.71% negativer Äußerungen von Mädchen über die Eigengruppe zu 39.53% negativer Angaben von Jungen über die Fremdgruppe deutlich unterschiedlicher und deuten eher auf eine positivere Bewertung/Wahrnehmung von Mädchen durch die Fremdgruppe der Jungen hin.

Wird für die Analyse der *schulbezogenen* Eigenschaftszuschreibungen das Geschlecht der ausfüllenden Person mit berücksichtigt, ergeben sich bei Angaben, welche durch Jungen gemacht wurden gleiche Ergebnisse im Vergleich zur Analyse ohne Berücksichtigung des Faktors Geschlecht (Aussagen über Jungen: 73 negative und 6 positive, schulbezogene Eigenschaften im Vergleich zu Angaben über Mädchen: 3 negative und 55 positive, schulbezogenen Eigenschaften), $\chi^2(1, N = 137) = 103.00, p < .001$. Auch bei durch Mädchen gemachten Angaben lässt sich das gefundene Muster replizieren – Jungen werden mit 92 negativen und 15 positiven Eigenschaften beschrieben, die Eigengruppe der Mädchen mit 11 negativen und 81 positiven, schulbezogenen Eigenschaften, $\chi^2(1, N = 199) = 108.56, p < .001$.

Im Rahmen einer Folgeanalyse wurden die gemachten Aussagen mit eindeutigem Schulbezug nicht mit der Gesamtzahl aller Aussagen kontrastiert, sondern nur mit dem Anteil *nicht* schulbezogener Aussagen einer weiteren Häufigkeitsanalyse unterzogen. Die Ergebnisse entsprechen in der Hauptsache dem Befundmuster der ersten Analyse und indizieren (wenngleich auch mit einem deutlich geringeren Unterschied), dass Jungen bei den nicht schulbezogenen Aussagen häufiger negativ und seltener positiv charakterisiert werden als Mädchen (Aussagen über Jungen: 66 positive und 261 negative; Aussagen über Mädchen: 101 positive, 197 negative; Aussagen $\chi^2(1, N = 625) = 14.97, p < .001$, vgl. Abbildung 4). Analysen getrennt nach Geschlecht zeigen erneut, dass Jungen ihre Eigengruppe als negativer beschreiben, als Mädchen (Angaben von Jungen über Jungen: 88 negative Aussagen, 27 positive Aussagen; über Mädchen: 65 negative Aussagen, 49 positive Aussagen, $\chi^2(1, N = 229) = 9.82, p < .01$). Bei Angaben von Mädchen zeigt sich dieses Befundmuster ebenso (Angaben von Mädchen über Jungen: 173 negative Aussagen, 39 positive Aussagen; über Mädchen: 132 negative Aussagen, 52 positive Aussagen, $\chi^2(1, N = 396) = 5.41, p < .05$). Die schulbezogenen Angaben bleiben hiervon unberührt, der dort gefundene Effekt zuungunsten von Jungen i.S.v. einer negativen Charakterisierung in der Schule im Kontrast zu Mädchen, die durchgehend nahezu ausschließlich mit positiven, schulischen Eigenschaften beschrieben werden bleibt sowohl ohne als auch mit Berücksichtigung des Faktors Geschlecht der Versuchsperson bestehen (statistische Angaben s.o.).

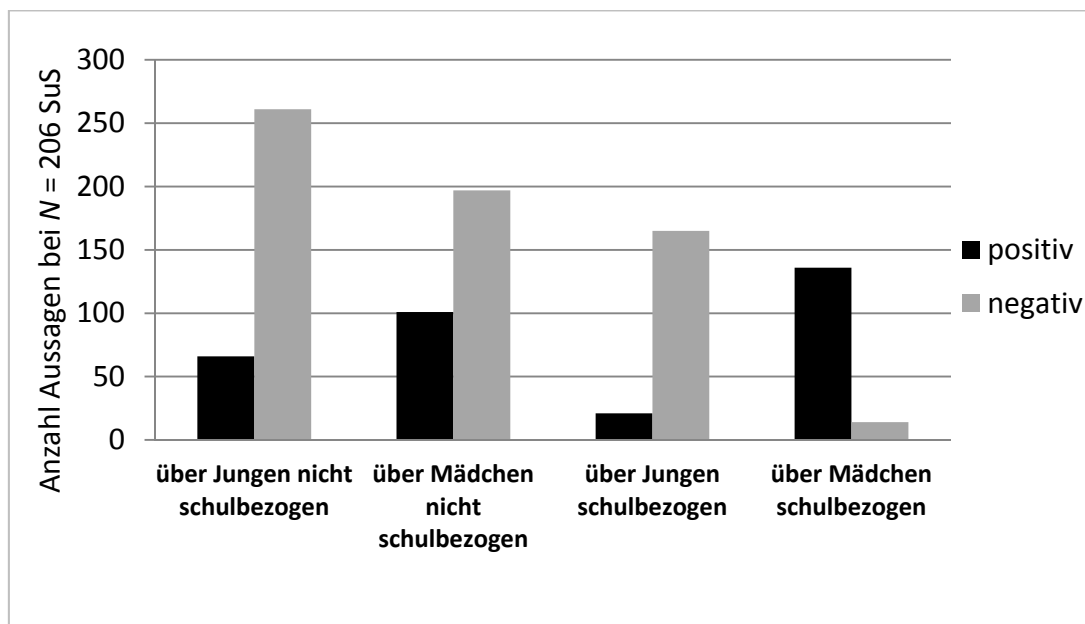


Abbildung 4: Positive und negative Aussagen allgemein sowie mit konkretem Schulbezug von $N = 206$ Schüler/-innen, nach Kontext (außerschulisch vs. schulbezogen) und Valenz (positiv vs. negativ).

Die geschlossenen Angaben aus den semantischen Differentialen wurden so kodiert, dass höhere Messwerte einer höheren Zustimmung zu *negativen* Verhaltenseigenschaften entsprechen. Die Mittelwerte der *Gesamtskala*, welche Jungen im Allgemeinen (Reliabilität: $\alpha = 0.91$) und Jungen in der Schule (Reliabilität: $\alpha = 0.95$) beschreiben, wurden danach einer 2 (Geschlecht

der Testperson) \times 2 (Kontext) *mixed model*-ANOVA mit dem Messwiederholungsfaktor „Kontext“ (im Allgemeinen vs. in der Schule) unterzogen.

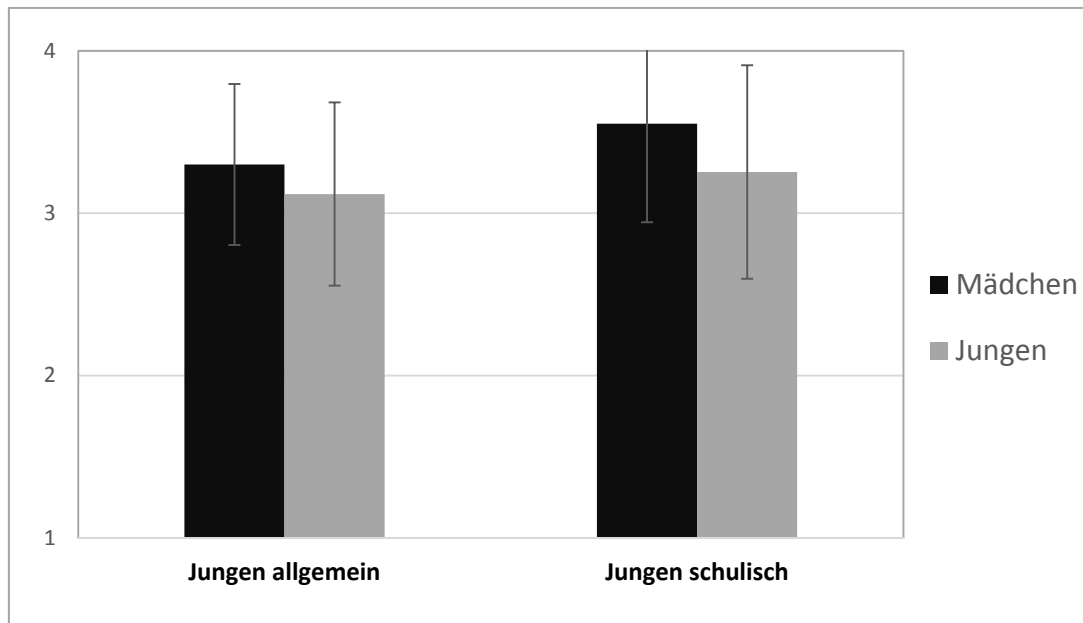


Abbildung 5: Mittelwerte auf den Subskalen der semantischen Differentiale über Eigenschaften von Jungen im Allgemeinen vs. in der Schule von Jungen nach Pb-Geschlecht. Schwarze Balken indizieren die Standardabweichung. Höhere Werte bezeichnen negativere Eigenschaftszuschreibungen.

Die Ergebnisse der Varianzanalyse zeigen einen signifikanten Haupteffekt für den Faktor Geschlecht, $F(1,200) = 11.67$, $p = .001$, $\eta^2 = .06$ in die Richtung, dass befragte Mädchen mit $M = 3.43$ signifikant negativere Beschreibungen von Jungen abgeben, als befragte Jungen dies tun, $M = 3.18$. Unabhängig vom Geschlecht der Testpersonen wurde ein Einfluss des Kontexts, auf welchen die Bewertung bezogen war, festgestellt, $F(1,200) = 22.31$, $p < .001$, $\eta^2 = .10$: Jungen wurden hierbei im Kontext Schule mit $M = 3.42$ negativer eingeschätzt, als mit $M = 3.21$ im Allgemeinen (Abbildung 5). Die Interaktion zwischen beiden Faktoren wurde mit $F < 1$ nicht statistisch bedeutsam.

5.4. Diskussion

Im medialen Diskurs werden, wie eingangs erläutert, Jungen als Verlierer im Schulsystem angesehen – Mädchen haben sie im Kontext Schule den medialen Darstellungen folgend seit Langem überholt. Das Ziel von Studie 1 war es zu überprüfen, inwiefern Schülerinnen und Schüler der Sekundarstufe durch diese für Jungen negative und für Mädchen positive Berichterstattung (Meta-)Stereotype erworben haben und damit einhergehend das Bild, was andere („die Leute“, also der gesellschaftliche Blick auf einer Metaebene) von Schülerinnen und Schülern haben mit schlechten schulischen Fähigkeiten seitens der Jungen und entsprechend negativem Verhalten im Kontext Schule assoziiert wird. In der freien Assoziationsaufgabe wurde festge-

stellt, dass deutlich mehr Äußerungen über negatives Verhalten von Jungen sowohl generell als auch in der Schule von beiden Geschlechtern produziert wurden. Analysen der *Ratingskala* mit verschiedenen semantischen Differentialen unterstützen diese Befunde: Hierbei wird Jungen in der Schule noch negativeres Verhalten unterstellt, als bereits im Allgemeinen.

Es konnte festgestellt werden, dass die befragten Jungen und Mädchen im Allgemeinen ein deutlich negativeres Bild zuungunsten der Jungen zeichneten, lediglich die Beschreibung der Schülerinnen und Schüler von Mädchen zeigte ein ausgeglichenes Bild i.S.v. nahezu gleich vielen positiven wie negativen Aussagen. Zum *ingroup bias* (Hypothesen 1 und 2) kann festgehalten werden, dass dieser sowohl bei der Beschreibung von Jungen als auch bei der von Mädchen nicht vorzuliegen scheint und diese Hypothesen abzulehnen sind. Während Jungen ein deutlich negativeres Bild von der Eigengruppe antizipierten, stellte sich die Verteilung von positiven und negativen Äußerungen bei Mädchen über die Eigengruppe zwar etwas ausgeglichener dar; von einer Bevorzugung der Eigengruppe im Sinne des *ingroup bias* kann allerdings hierbei nicht ausgegangen werden. Auch wenn die Charakterisierung von Eigen- und Fremdgruppe auf einer Metaebene erfolgte, wäre nichtsdestotrotz ein gewisses Ausmaß an Eigengruppenfavorisierung zu erwarten, da Kognitionen über die eigene soziale Gruppe angestoßen werden und somit die soziale Identität zumindest in gewissem Ausmaß aktiviert wird.

Im Hinblick auf die Debatte über Jungen als Bildungsverlierer stellte sich die letzten Annahme (Hypothese 3) als besonders bedeutsam heraus, da hier erstmals der konkrete Schulbezug der Äußerungen in die Analysen mit aufgenommen wurde. Es konnte gezeigt werden, dass von den befragten Schülerinnen und Schülern deutlich mehr negative Aussagen über Jungen mit einem klaren Bezug zum Kontext Schule produziert wurden, als über Mädchen. Weiterhin zeigten die Analysen, dass dieses Muster nicht vom Geschlecht der Probanden/-innen abzuhängen scheint und bei befragten Jungen nahezu äquivalent zu den Angaben von Mädchen auftritt, was den erwarteten Effekten entspricht.

Zusammenfassend kann festgehalten werden, dass Jungen durch die befragten Schüler/-innen in der Hauptsache mit negativen Persönlichkeits- oder Verhaltenseigenschaften charakterisiert wurden, während Mädchen überwiegend positiv dargestellt wurden. Dieses Befundmuster betrifft sowohl allgemeine als auch (in verstärktem Ausmaß) schulbezogene Äußerungen beider Geschlechtergruppen, ein *ingroup bias* scheint nicht vorzuliegen. Die Beschreibungen der Schüler/-innen ergeben somit deutliche Evidenz für das Vorliegen von (Meta)-Stereotypen ausschließlich zuungunsten von Jungen und entsprechend der medialen Berichterstattung über deren Versagen in der Schule verstärkt in ebendiesem Kontext. Auch bei getrennten Analysen hinsichtlich des Geschlechts der ausfüllenden Person konnten keine Abweichungen von diesem Befundmuster festgestellt werden – die Annahmen scheinen Jungen wie Mädchen

gleichsam akquiriert zu haben. Zu berücksichtigen bei der Interpretation der Ergebnisse bleibt, dass in diesem Teil der Untersuchung auf indirekte Art das Vorhandensein stereotyper Annahmen überprüft wurde (genauer siehe Limitationen).

Die Analysen der semantischen Differentiale unterstützen und ergänzen die gefundenen Effekte der offenen Befragung: Auch hier wurden Jungen von beiden Geschlechtern im Kontext Schule deutlich negativer beschrieben, als im Allgemeinen. Zusätzlich haben die befragten Schülerinnen hier unabhängig vom Kontext (Allgemein vs. Schule) eine negativere Verhaltensbeschreibung von Jungen abgegeben, als es bei den untersuchten Schülern der Fall war. Diese Ergebnisse zeigen ebenfalls, dass die befragten Schüler/-innen die Ansicht vertreten, dass das gesellschaftliche Bild von Jungen eher von negativen Erwartungen – insbesondere in der Schule – geprägt ist. Die hier gefundenen Ergebnisse stimmen insgesamt mit den Befunden an einer jüngeren Stichprobe in Großbritannien von Hartley & Sutton (2013) überein.

Nach der *Social Identity Theory* (SIT; Tajfel & Turner, 1986 – vgl. ausführlicher Kap. 4) strebt jedes Individuum und damit auch jede soziale Gruppe nach einer positiven Selbsteinschätzung. Zu einer solchen Einschätzung gelangen Personen über die Evaluation ihrer Gruppenmitgliedschaft und der damit verbundenen Bewertung der Gruppe. Diese wiederum speist sich aus Intergruppenvergleichen, bei welchen stets die Herstellung einer positiven Distinktheit von einer relevanten *Outgroup* Ziel sein sollte. Ist der Vergleich hingegen negativ, bieten sich diverse Möglichkeiten an. Es ist denkbar, a) die eigene Gruppe zu verlassen, b) die eigene Gruppe aufzuwerten bzw. c) sich mit den wahrgenommenen Unterschieden abzufinden und den niedrigeren Status der *Ingroup* zu akzeptieren und (um die positive Distinktheit wiederherzustellen) andere Vergleichsdimensionen heranzuziehen bzw. eine andere Kategorie jenseits von Geschlecht zur gruppendifinierenden Variable zu erklären. Da beim sozialen Kategorisierungsfaktor „Geschlecht“ die soziale Mobilität gegen Null tendiert und Geschlecht im Jugendalter eine herausragend wichtige Kategorie für die eigene Identität darstellt (Hannover & Greve, 2012), verbleibt für die Jungen im vorliegenden Experiment nur die Möglichkeit, die eigene Gruppe aufzuwerten oder die Vergleichsdimension zu ändern. Dies kann positive Konsequenzen für Jungen haben, sich aber auch sehr negativ für Jungen sowie die Intergruppenbeziehung zwischen beiden Geschlechtern auswirken. In einer Erweiterung zur SIT, der Selbstkategorisierungstheorie (SKT) von Turner, Hogg, Oakes, Reicher und Wetherell (1987) wird angeführt, dass bei Salienz der Vergleichsdimension (hier u.a.: schulisches Verhalten und Leistungsfähigkeit) ein Mechanismus der Depersonalisation auftritt. Ein *Priming* auf die soziale Gruppe, der ein Individuum angehört, führt hierbei der Theorie folgend zu einem Wechsel der aktuell relevanten Identitätsdimension, weg von der Fokussierung auf eine individuelle Identität hin zur Einnahme einer sozialen Identität. Bei hoher wahrgenommener Ähnlichkeit zwischen Individuum und Gruppe führt dies dazu, dass eine Person auf Stereotypen

über die eigene Gruppe (wie in der vorliegenden Studie gemessen) zurückgreift und damit konsistentes Verhalten zeigt (*self-stereotyping*). Unter Integration der Ergebnisse aus Studie 1 kann somit vermutet werden, dass die durch die Schüler/-innen angegebenen kognitiven Repräsentationen über das Verhalten von Jungen und Mädchen insbesondere in der Schule keineswegs nur die bloße Auflistung von potentiell Verhalten darstellt, sondern dieses Verhalten dann auch in entsprechenden Situationen tatsächlich so gezeigt wird. Dies könnte eine wichtige Erklärungsmöglichkeit für die in der Einleitung dargelegten Unterschiede hinsichtlich schulischen Verhaltens von Mädchen und Jungen bereitstellen. Dass wahrgenommene Stereotype – insbesondere bei hoher Identifikation mit der Eigengruppe – dann durch eigenes Verhalten bestätigt werden, wurde in mehreren empirischen Studien belegt (Latrofa, Vaes & Cadinu, 2012; Latrofa, Vaes, Cadinu & Carnaghi, 2010; Spears et al., 1997, im schulischen Kontext: Chatard, Guimond & Selimbegovic, 2007). Auch nicht direkt verhaltens- oder leistungsbezogene Maße werden hierdurch beeinträchtigt, wie z.B. Attributionen von Erfolg oder Misserfolg (Nicholls, 1975; Stipek & Gralinski, 1991), die Intergruppenbeziehung in Richtung eines verstärkten negativen Affekts bei Antizipation eines Kontaktes mit einem Fremdgruppenmitglied (Vorauer et al., 1998), gesteigerte Aggression und Delinquenz (Issmer, Stellmacher & Gollwitzer, 2013), negative selbstbezogene Gefühle (Vorauer & Kumhyr, 2001) sowie negative Auswirkungen auf den Selbstwert einhergehend mit einer verminderten Selbstkonzeptklarheit von Personen der *Ingroup* (Vorauer et al., 1998).

Diese Befunde indizieren möglicherweise weitreichende Konsequenzen für die gefundenen Ergebnisse. Getriggert durch die negative Wahrnehmung von Jungen ist einerseits denkbar, dass diese aufgrund fehlender, alternativer Handlungsmöglichkeiten nach SIT und SKT durch die Antizipation negativer Erwartungen entsprechende negative Verhaltensweisen auch bewusst zeigen, um sich der eigenen Geschlechtsidentität rückzuersichern – mit fatalen Konsequenzen für deren schulischen Werdegang. Andererseits werden neben möglichen Auswirkungen i.S.v. *self-stereotyping* weitere negativen Konsequenzen immanent, die mit schulischem Erfolg nicht vereinbar sind (geringeres schulisches Selbstkonzept, Aggression, etc.).

Abschließend stellen aktuelle Forschungsarbeiten eine zusätzliche Erklärung für das Fehlen eines *ingroup-bias* in der vorliegenden Studie bereit. Einerseits argumentieren Latrofa et al. (2010), dass negative Wahrnehmungen über die Eigengruppe dazu führen können, dass (positive, wie auch negative) Aspekte des wahrgenommenen Stereotyps ein Bestandteil des Selbst werden. Da aufgrund der Ergebnisse der hier präsentierten Studie davon auszugehen ist, dass Stereotype über Jungen im Allgemeinen und in der Schule in der Hauptsache negativ besetzt sind, kann dies u.U. bereits dazu geführt haben, dass Jungen selbst keine positiven Aspekte mehr wahrnehmen und die eigene Geschlechtsidentität somit hauptsächlich über negativ konnotierte Eigenschaften konstruieren (Jones & Myhill, 2004). Ferner ist abweichendes Verhalten von Jungen im Jugendalter durchaus positiv besetzt, d.h. möglicherweise ist es für

Jungen instrumentell, sich so zu beschreiben und sich dann auch so zu verhalten. Demnach ist die Möglichkeit gegeben, dass Jungen durch negatives Verhalten im Allgemeinen, insbesondere aber im *feminin* konnotierten Kontext Schule ihre Geschlechtsidentität stabilisieren, indem sie schulisch *unangepasstes* Verhalten zeigen (Hannover & Kessels, 2011).

Ergänzend weisen Befunde von Verkuyten & Nekuee (1999) darauf hin, dass *self-stereotyping* eine eigentlich für die Eigengruppe sprechende Verzerrung auf negative Inhalte der antizipierten Stereotype umlenken kann. Dieser Befund erweist sich allerdings nur für Personen gültig, die sich stark mit der Eigengruppe identifizieren. Aufgrund der mehrfach dargelegten Relevanz der Kategorie „Geschlecht“ im Jugendalter ist allerdings in der vorliegenden Untersuchung von einem hohen Anteil *high identifiers* auszugehen, was für die Übertragbarkeit der Befunde auf den vorliegenden Kontext spricht.

Limitationen. Eine Besonderheit der vorliegenden Untersuchung ist der Ansatz nicht die eigene Meinung von Schülerinnen und Schülern zu erfragen, sondern die Frage auf Dritte („die Leute“) zu beziehen. Zwar ist es zur Beantwortung der Untersuchungsfragestellung dringend geboten, dass die Angaben der Schüler/-innen nicht (oder nicht allzu stark) durch ihre eigene Meinung konterkariert werden, da es in dieser Untersuchung vielmehr um die Feststellung geht, ob die befragten Schüler/-innen die Ansicht vertreten, dass das über die Medien transportierte Bild tatsächlich die gesellschaftliche Meinung und Sichtweise widerspiegelt. Allerdings stellt es sich für die befragten Personen u.U. als schwierig dar, sich in eine entsprechende Situation hineinzusetzen. Durch eine sorgfältige Instruktion mit einem konkreten Beispiel wurde versucht, mögliche Schwierigkeiten abzubauen. Im Hinblick auf die nicht haltbaren Hypothesen zum *ingroup bias* scheint es gelungen zu sein, die Schüler/-innen zum Eindringen in die gesellschaftliche Meinung zu bewegen und die eigenen Ansichten dabei weitestgehend unberücksichtigt zu lassen. Zwar bleiben auch andere Erklärungen für das Nichtauffinden des *ingroup bias* zu berücksichtigen, z.B. dass dieser bei Frauen häufig stärker auftritt, als bei Männern (Lindeman, 1997; Rudman & Goodwin, 2004). Da jedoch keinerlei Anhaltspunkt auf eine Verzerrung zugunsten der Eigengruppe – bei beiden Geschlechtern – festgestellt werden konnte, diese aber entsprechend der SIT (Tajfel & Turner, 1986; vgl. auch Lindeman, 1997; Turner & Reynolds, 2001) als Mechanismus zur Absicherung einer positiven Distinktheit der *Ingroup* zumindest ansatzweise zu erwarten gewesen wäre, ist dieses Ergebnis möglicherweise ein Indikator für ein gelungenes Instruktionsverständnis. Es bleibt dennoch zu berücksichtigen, dass sich in den gemachten, offenen Angaben immer auch anteilig individuelle Meinungen und Ansichten der befragten Schüler/-innen widerspiegeln, wobei nicht vollständig aufgeklärt werden kann, wie stark die gemachten Angaben individuell verzerrt wurden. Weiterhin ist die eingenommene, gesamtgesellschaftliche Perspektive und Fokussierung auf „die Leute“ trotz ihrer Vorzüge für diese Untersuchung nicht vollkommen unproblematisch, da die Einnahme

einer derartigen Position eine nicht unerhebliche kognitive Leistung erfordert. Neben den eigenen Ansichten müssen konsequent andere Sichtweisen und Meinungsbilder antizipiert werden, zusätzlich muss es gelingen eine ggf. abweichende, eigene Meinung zumindest in einem gewissen Maße zurückzustellen. Diese kognitiven Ressourcen stehen wiederum für die Aufgabenbearbeitung nicht mehr zur Verfügung. Gleichfalls unklar bleibt, inwiefern die Untersuchungsergebnisse mit einer impliziteren Messung repliziert werden könnten.

Ebenso bleibt zu berücksichtigen, wie stark die Probanden/-innen selbst an das Vorhandensein entsprechender Stereotype glauben, da die niedergeschriebenen Eigenschaften nicht das Resultat einer individuellen, sondern vielmehr einer antizipierten kollektiven Meinungsbildung darstellen. Auf die Wichtigkeit des Glaubens an ein bestimmtes Stereotyp und dessen Akzeptanz weisen u.a. Chatard, Selimbegovic, Konan & Mugny (2008), Leyens et al., (2000) sowie Schmader et al., (2004) hin. Eine entsprechende Erweiterung der Befragung um Maße der individuellen Zustimmung zu geäußerten Stereotypen könnte zukünftig mit einer noch höheren Aussagekraft der Befunde einhergehen. Daher könnte eine erneute Befragung der *individuellen* Wahrnehmung über Jungen im Allgemeinen und in der Schule („was denkst *du*“) als interessante Studie zur Kontrastierung dienen.

Bei den offenen Angaben gab es zusätzlich Schwierigkeiten bei der Kodierung der freien Äußerungen, insbesondere bei der Feststellung, ob eine Verhaltenseigenschaft (z.B. „fleißig“) einen deutlichen Schulbezug aufweist oder nicht. Durch eine gewissenhafte Schulung der Kodierer und die Sicherstellung einer vollkommen unabhängigen Kodierung (ohne Möglichkeiten zur Absprache zwischen den Kodierenden) konnte jedoch eine zufriedenstellende *Intercoder*-Reliabilität erzielt werden, was für das Gelingen der Klassifikation spricht. Für zukünftige Arbeiten wäre denkbar, *post-hoc* zumindest die am häufigsten genannten Begriffe von Lehrer/-innen oder Lehramts-Studierenden hinsichtlich des Schulbezugs bewerten zu lassen und die daraus resultierenden Ergebnisse der Kodierung zugrunde zu legen bzw. gemachte Aussagen mit der *post-hoc* erhobenen Stärke des Schulbezugs zu gewichten

Die geschlossenen Angaben (semantische Differentiale) waren – aufgrund zeitlicher Restriktionen bei der Untersuchung – ausschließlich auf das Verhalten von Jungen im Allgemeinen und in der Schule bezogen. Die Sichtweise auf Mädchen in beiden Kontexten blieb hierbei leider unberücksichtigt. Für eine bessere Kontrastierung und einen genaueren Vergleich beider Gruppen wäre es daher wertvoll, zusätzlich auch die Gruppe der Mädchen mit einzubeziehen und diese entsprechend als weiteren Faktor in die Varianzanalyse aufzunehmen. Trotz partieller η^2 von .06 bis .10 und damit einer mittleren Effektgröße, lässt sich diese vermutlich durch die Ergänzung von Angaben über Mädchen in beiden Kontexten noch bedeutsam steigern.

Abschließend sei auf mögliche Herkunftseffekte verwiesen, welche in den vorliegenden Ergebnissen nur in eingeschränkter Art und Weise kontrolliert werden können. Die ethnische

Herkunft ist nur über das eigene Geburtsland der Schüler/-innen nachzuvollziehen, da aufgrund einer behördlichen Anordnung keinerlei Daten über die Eltern erhoben werden konnten. Der über die Sprache gebildete Indikator kann zwar als Schätzer herangezogen werden, ist allerdings wesentlich ungenauer. Kongruent liegen für die soziale Herkunft nur wenige Indikatorvariablen vor. Da sich die befragten Gruppen von Jungen und Mädchen auf den gebildeten Indikatoren nicht unterschieden, kann davon ausgegangen werden dass die berichteten Befunde nicht auf Grundlage der sozialen oder ethnischen Herkunft, sondern wie intendiert nur auf Grundlage des eigenen Geschlechts zustande gekommen sind. Befunde von Segeritz et al. (2010) weisen ebenfalls darauf hin, dass der ethnischen Herkunft nur unter bestimmten Voraussetzungen eine Rolle beim Zustandekommen von Unterschieden in Leistung und anderen, schulbezogenen Variablen zukommt, so z.B. scheint eine türkische Herkunft das Risiko von Jungen für einen geringeren Bildungserfolg und Schulentfremdung zu verstärken, was es wünschenswert erscheinen lässt, in zukünftigen Untersuchungen Variablen zur (ethnischen und sozialen) Herkunft genauer zu erfassen, um potentielle Einflüsse dieser Merkmale auf die Ergebnisse genauer analysierbar zu machen.

Trotz der beschriebenen Einschränkungen liefert diese Studie wertvolle Grundlagendaten und zeichnet ein aktuelles Bild über die Wahrnehmung von Jungen in der Schule im Kontext der medialen Berichterstattung über diese soziale Gruppe als generelle Schulversager. Die Ergebnisse dieser Studie suggerieren, dass Schüler/-innen bereits negative Erwartungen über das Verhalten von Jungen internalisiert haben oder – anders formuliert – durch die Medien dargestellte Verhaltensweisen von Jungen (in der Schule) selbst glauben und antizipieren.

Um festzustellen, welche Auswirkungen eine Exposition bzw. Konfrontation mit den entsprechenden (Meta-)Stereotypen – vermittelt über die mediale Darstellung und Verbreitung des Bildes des schulversagenden Jungen – auf verschiedene Variablen wie Leistung und Motivation der Schülerinnen und Schüler haben kann, wurden die im Folgenden vorgestellten Studien konzipiert.

6. Studie 2: Stereotype Threat und Leistung bei Jungen

Auswirkung der medialen Berichterstattung auf die Leistungsfähigkeit von Jungen

Veröffentlicht als Studie 2 in:

Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!?! How the public discourse on failing boys impacts performance outcomes and motivational goal orientation in German students. *Social Psychology, 45*, 112-126.

6.1. Überblick und Fragestellung

Wie im Theorieteil ausführlich dargelegt wurde, ist *stereotype threat*, also die Minderleistung infolge einer Konfrontation eines negativen Leistungsstereotyps und der daraus resultierenden Befürchtung, dieses Stereotyp zu bestätigen, ein vielfach untersuchtes Konstrukt in der Sozialpsychologie. Während sich vorangegangene Forschung jedoch verstärkt auf die Benachteiligung von Mädchen und Frauen in mathematisch-naturwissenschaftlichen Bereichen (MINT) konzentrierte (Weaver-Hightower, 2003; Nguyen & Ryan, 2008), ist bisher weitgehend ungeklärt, inwieweit der Paradigmenwechsel und die Fokussierung der medialen Berichterstattung auf die nachlassende Leistung von Jungen zur Ausbildung von negativen Leistungsstereotypen bei Jungen geführt hat und inwieweit diese Stereotypen im Sinne einer Bedrohung leistungsmindernde Wirkungen entfalten. Studie 1 liefert dahingehend Evidenz, dass die Darstellung von Jungen als „Schulversager“ durch die mediale Verbreitung und womöglich über entsprechende Multiplikatoren (z.B. Eltern oder Lehrer) dazu geführt hat, dass Jungen weniger positiv wahrgenommen werden (und Jungen die Eigengruppe auch negativer beschreiben) sowie insbesondere in der Schule verstärkt und beinahe ausschließlich mit negativen Lern- und Leistungseigenschaften assoziiert werden. Demzufolge kann angenommen werden, dass das öffentlich diskutierte Stereotyp über bildungsschwache Jungen auch bei den untersuchten Schüler/-innen angekommen ist und internalisiert wurde. In dieser Studie soll nun geklärt werden, ob und wie sich diese negativen Leistungserwartungen für Jungen im Sinne von *stereotype threat* negativ auf die Leistung, gemessen über standardisierte Schulleistungstests in den zwei schulischen Hauptfächern Deutsch und Mathematik, auswirken können.

Die Befunde aus der bisherigen *stereotype threat*-Forschung zu Frauen im MINT-Bereich sollen somit einerseits auf Jungen, andererseits auf nicht nur auf einen spezifischen Leistungsbe- reich, in welchem ein negatives Leistungsstereotyp über eine soziale Gruppe besteht übertragen werden, sondern auf ein breiteres Spektrum von Schulleistungen. Bisherige Forschungsergebnisse zeigen, dass ein Großteil der durchgeführten Untersuchungen zu Frauen in Mathematik nachweisen konnte, dass eine Aktivierung des entsprechenden Stereotyps zu einer signifikant verminderten Leistung in der bedrohten Domäne führt. Es liegen allerdings auch Studien vor, die keine derartigen Effekte nachweisen können (Nguyen & Ryan, 2008). Zu berücksichtigen bleibt auch, dass die Übertreibung einer Bedrohungssituation einen gegenteiligen Effekt haben kann (Kray et al., 2001, 2004) und eine Reaktanz verbunden mit einer vergleichsweise besseren Leistung entsteht. Dies macht deutlich, dass die Ableitung von konkreten Hypothesen zwar möglich ist, dass dies aber durch die durchaus divergierenden Ergebnisse (Stoet & Geary 2012) und die bisher kaum vorhandenen Studien über die Auswirkungen bei einer Stereotypenbedrohung bei Jungen erschwert wird. Erste Hinweise auf eine Übertragbarkeit bisheriger Ergebnisse zu *stereotype threat* auf die Jungenforschung lieferten Hartley & Sutton (2013), welche nachwiesen

dass eine generelle Stereotypenbedrohung von Jungen zu umfassenden Leistungseinbußen in allen erhobenen Domänen (Lesen, Schreiben und Mathematik) führen. Konträr dazu argumentierten Eckert & Imhof (2013) auf Basis ihrer Ergebnisse, dass die schädlichen Auswirkungen von *stereotype threat* bei Mädchen nicht auf Jungen übertragbar sind und zeigten, dass nach einer Bedrohung von Jungen in einem Lesetest keinerlei bedeutsame Unterschiede auftraten.

Trotz der Probleme, die aus dem divergierenden Forschungsstand erwachsen, sollen folgende Hypothesen zur Prüfung im Rahmen dieser Studie formuliert werden:

- H4* Jungen in der Experimentalbedingung (*stereotype threat*), welche eine generelle Bedrohung ihrer schulischen Leistung erfahren, zeigen signifikant verminderte Leistungen in Deutsch und Mathematik, im Sinne eines Leistungsabfalls vom Pretest zum Posttest im Vergleich zu Jungen der Kontrollgruppe.
- H5* Mädchen in der Experimentalbedingung verbessern ihre Leistung im Vergleich zu Mädchen der Kontrollgruppe, da sie nicht Mitglieder der stereotypisierten Gruppen sind und Leistungsvorteile antizipieren (*stereotype lift*).

6.2. Methode

Teilnehmende. Für die Studie wurden 124 Schüler/-innen (davon 69 Jungen, 55 Mädchen) einer Berliner Sekundarschule (Realschule) des Berliner Stadtteils Steglitz untersucht. 20 Jungen und 4 Mädchen wurden aus der weiterführenden Analyse ausgeschlossen, da diese nicht ernsthaft an der Untersuchung teilnahmen und insbesondere das experimentelle Treatment einschließlich der sehr einfachen Kontrollfragen (siehe Prozedur) boykottierten (unter 10% der Leistungsaufgaben bearbeitet, nur 25% oder weniger der Kontrollfragen zum Treatment korrekt beantwortet). Auf die verbleibenden $N = 100$ Schüler/-innen entfielen nunmehr 49 Jungen und 51 Mädchen, welche im Durchschnitt $M = 14.61$ Jahre ($SD = .70$; $range = 14-17$ Jahre) alt waren. Hinsichtlich des Migrationshintergrundes hatten $n = 4$ Jungen und $n = 3$ Mädchen eine direkte Migrationserfahrung, der Rest der Stichprobe gab an, in Deutschland geboren zu sein. Dies entspricht einer geringen Quote von Migranten/-innen im Klassenverband (vgl. Stanat et al., 2010), spiegelt aber realistisch den geringen Anteil der Migranten im Schulbezirk wieder. Angaben über den Migrationshintergrund der Eltern durften nach Vorgabe durch die genehmigende Schulaufsichtsbehörde nicht erfasst werden, daher wurde diesbezüglich ein Indikator über die mit den Eltern gesprochene Sprache gebildet. Hierzu liegen Angaben von $n = 82$ Schülerinnen und Schülern vor. Von insgesamt 37 Angaben, die Jungen machten, sprachen $n = 12$ (32.4%) eine andere Sprache mit den Eltern, was auf einen Zuwanderungshintergrund in der zweiten Generation deutet. Von 45 Angaben, die von Mädchen zur Sprache mit den Eltern vorliegen sprachen $n = 13$ (28.9%) eine andere Sprache – die gesprochene Sprache mit den Eltern unterscheidet sich we-

der zwischen den Geschlechtern, noch zwischen den Untersuchungsbedingungen, $\chi^2 < 1, p > .70$ was für eine Gleichverteilung der gesprochenen Sprache resp. der indirekten Migrationserfahrung zwischen den jeweiligen Gruppen spricht.

Zur Erfassung des sozioökonomischen Status wurden die gleichen PISA-Items zu „*home possessions*“ eingesetzt, die auch in Studie 1 zur Anwendung kamen. Es zeigte sich eine mittlere Verfügbarkeit der abgefragten Güter (u.a. Bücher für die Schule, Schreibtisch zum Lernen und Arbeiten) von $M = 5.61$ ($SD = 0.96$, $range = 3-7$) verfügbaren Gütern, was einem gehobenem sozioökonomischen Status entspricht (OECD, 2011). Eine univariate Varianzanalyse machte deutlich, dass Probanden/-innen der Kontrollgruppe eine leicht höhere Verfügbarkeit der abgefragten Güter angaben ($M = 5.76$, $SD = 0.87$), als Teilnehmende in der Experimentalgruppe ($M = 5.46$, $SD = 1.03$), $F(1,96) = 2.83$, $p = .10$, $\eta^2 = .03$, wobei beide Mittelwerte jedoch einem eher hohen sozioökonomischen Status zuzuordnen sind. Weder der Haupteffekt für Geschlecht ($M_{\text{Jungen}} = 5.71$, $SD = 0.89$; $M_{\text{Mädchen}} = 5.51$, $SD = 1.03$), noch der Interaktionsterm indizierte weitere Unterschiede hinsichtlich dieses Merkmals, $F < 1.51$, $p > .22$.

Zusätzlich zum Migrationshintergrund und dem sozioökonomischen Status wurde die aktuelle Schulnote in Deutsch und Mathematik erfragt, um ein initiales Leistungsmaß zu erfassen. Mit einer durchschnittlichen, aktuellen Deutschnote von $M = 2.99$ ($SD = .80$; $range = 1-5$) und einer Mathematiknote von $M = 3.13$ ($SD = 1.07$; $range = 1-5$) zeigte sich auf beiden Variablen ein sehr breites Noten- und Leistungsspektrum von sehr gut bis ungenügend, welches im Mittel auf ein befriedigendes, mittleres Niveau in beiden Fächern hinweist.

Um kontrollieren zu können, ob sich die Deutsch- und Mathematikleistung/-note zwischen beiden Geschlechtern bedeutsam unterscheidet, wurden Einzelvergleiche in Abhängigkeit des Geschlechts berechnet. Die folgende Tabelle gibt eine Übersicht über die Ergebnisse:

Tabelle 4: Einzelvergleiche aktueller Deutsch- und Mathematiknote in Abhängigkeit des Geschlechts der Studienteilnehmer/-innen. Zu $N = 100$, resp. $df = 98$ fehlende Fälle entsprechen der Anzahl fehlender Angaben.

Variable	Mittelwerte	T-Wert (df)	p
	Geschlecht		
aktuelle Deutschnote	Jungen $M = 2.96$	-0.44 (96)	.66
	Mädchen $M = 3.03$		
aktuelle Mathematiknote	Jungen $M = 3.16$	0.27 (95)	.79
	Mädchen $M = 3.10$		

Die Schüler/-innen wurden nun randomisiert auf die Untersuchungsbedingung zugewiesen, es entfielen auf die Experimentalgruppe (*stereotype threat*) insgesamt $n = 27$ Jungen und $n = 23$ Mädchen; die Kontrollgruppe (*no threat*) bestand aus $n = 22$ Jungen und $n = 28$ Mädchen. Erneut wurden zur Überprüfung der Homogenität beider Gruppen auf den Leistungsmaßen (Schulnoten) in Deutsch und Mathematik Einzelvergleiche angestellt, eine Übersicht über die Ergebnisse gibt Tabelle 5:

Tabelle 5: Einzelvergleiche aktueller Deutsch- und Mathematiknote in Abhängigkeit der Untersuchungsbedingung. Zu $N = 100$, resp. $df = 98$ fehlende Fälle entsprechen der Anzahl fehlender Angaben.

Variable	Mittelwerte	T-Wert (df)	p
	Untersuchungsbed.		
aktuelle Deutschnote	EG $M = 3.07$	0.95 (96)	.35
	EG $M = 2.92$		
aktuelle Mathematiknote	EG $M = 3.11$	-0.15 (95)	.88
	KG $M = 3.15$		

Zusammengefasst weisen hinsichtlich der soeben beschriebenen Hintergrundmerkmale sowie der ebenfalls erfragten, aktuellen Deutsch- und Mathematiknote die angestellten Analysen zwischen den Gruppen (Geschlecht und Untersuchungsbedingung) mit Ausnahme des *tendenziellen* Befundes hinsichtlich des sozioökonomischen Status zwischen den Untersuchungsbedingungen keine bedeutsamen Unterschiede zwischen den Geschlechtern oder innerhalb der Untersuchungsbedingung auf, was für eine Vergleichbarkeit beider Geschlechter und beider Untersuchungsgruppen als notwendige Grundlage für die geplante, experimentelle Manipulation spricht.

Prozedur und Material. Die Erhebung fand in zwei aufeinander aufbauenden, regulären Schulstunden am Vormittag im gewohnten Klassenverband statt. Um den Einfluss des initialen Leistungsstandes bei der Erhebung zu kontrollieren wurde ein Pre-Posttest-Design gewählt. Zuerst erfolgte eine kurze mündliche Instruktion, welche die Schüler/-innen mit dem zu bearbeitenden Testheft und den Aufgabenformaten vertraut machte. Ergänzend wurde zur Erhöhung der *Compliance* und zum Ansporn, die Testhefte gewissenhaft und vollständig zu bearbeiten, eine Entlohnung von 10 Euro je Fragebogen an die Klassenkasse zugesagt. Danach

wurden die Testhefte in randomisierter Reihenfolge und unter Zusicherung absoluter Anonymität verteilt. Zunächst war der Prä-Leistungstest zu bearbeiten, welcher aus 15 curriculumsnahen Aufgaben zum Leseverständnis und 12 Aufgaben aus der Mathematik bestand. Welches Fach zuerst zu bearbeiten war, wurde innerhalb der Stichprobe ausbalanciert. Die Fragen des Leistungstests wurden unveröffentlichten Aufgaben aus den Vergleichsarbeiten 2008 (Institut zur Qualitätsentwicklung im Bildungswesen, 2008) entnommen und bildeten ein breites Spektrum an Sub-Kompetenzbereichen (u.a. Schlussfolgern, Begründen in Deutsch, Gleichungen, Geometrie in der Mathematik) ab. Die Aufgabenschwierigkeit der vorgetesteten Aufgaben war zum Großteil im mittleren Niveau angesiedelt, jeweils zwei Aufgaben waren jedoch deutlich leichter (90% und mehr Lösungswahrscheinlichkeit) bzw. schwieriger (20% und weniger Lösungswahrscheinlichkeit). Die Beantwortung der Aufgaben erfolgte entweder mittels Auswahl der richtigen Antworten aus einer Liste vorgegebener Antwortmöglichkeiten (*multiple choice*) oder durch eine kurze, offene Beantwortung mit wenigen Worten. Für die Beantwortung jedes Aufgabenteils (Deutsch / Mathematik) waren je 15 Minuten Bearbeitungszeit vorgesehen.

Nach einer kurzen Pause von 10 Minuten zur Erholung der Versuchsteilnehmer/-innen wurde das experimentelle Treatment appliziert. In der Experimentalgruppe lasen die Probandinnen und Probanden einen kontinuierlichen Text von ca. einer Seite Länge, welcher aus dem Magazin *Der Spiegel* entnommen, gekürzt und sprachlich vereinfacht wurde. Dieser Artikel (Quelle: „Schlaue Mädchen, dumme Jungen“, *Der Spiegel*, 17.05.2004) wurde hierbei als interdisziplinäre Leistungsaufgabe maskiert. Es wurde ausgeführt, dass Jungen in der Schule generell weniger Kompetenzen erwerben und vergleichsweise geringere Schulformen besuchen als Mädchen. Dies wirke sich wiederum nicht nur auf die Qualität des Schulabschlusses aus, sondern habe auch gesamtgesellschaftliche Relevanz für die Zukunft von Jungen und jungen Männern. Es wurden keinerlei Informationen über domänen- bzw. fachspezifische Unterschiede zwischen den Geschlechtern gegeben. Zusätzlich wurden die Ausführungen aus dem Text mit einer Abbildung unterstützt, welche die Schulabschlüsse getrennt nach Geschlecht darstellten und aus welcher ersichtlich wurde, dass mit steigender Schulform der Anteil von Jungen abnimmt. Die Kontrollgruppe bearbeitete einen Text gleicher Struktur und Länge, welcher (geschlechtsneutral formuliert) Veränderungen im deutschen Schulsystem darlegte (Quelle: „Warum sich ein deutsches Bildungssystem neu erfinden muss“, *Der Spiegel*, 09.11.2009).

Zusätzlich wurden zum gelesenen Text vier Kontrollfragen gestellt (2 offene Fragen, 2 *single choice*-Fragen), um sicherzustellen, dass die Versuchsteilnehmer/-innen alle relevanten Informationen aus dem Text und der Abbildung abgeleitet haben. Für das Lesen des Textes zur Induktion von *stereotype threat* vs. der Kontrollbedingung waren 10 Minuten vorgesehen.

Aufbauend auf das experimentelle Treatment erfolgte der Posttest, welcher aus ebenfalls 15 Aufgaben zum Leseverständnis und 12 mathematischen Aufgaben bestand (Reihenfolge

erneut ausbalanciert). Die Auswahl der Aufgaben erfolgte dabei so, dass sowohl Kompetenzbereich als auch Aufgabenschwierigkeit mit den Anforderungen des Pretests übereinstimmten, die Bearbeitungszeit betrug erneut 15 Minuten je Aufgabenbereich.

Den Abschluss bildete ein demografischer Fragebogen, in welchem Alter, Geschlecht und Herkunft (Geburtsland: Deutschland vs. anderes Land) der Teilnehmer/-innen, häusliche Besitztümer (*home possessions*, Frey et al., 2009), gesprochene Sprachen (z.B. mit den Eltern: Deutsch vs. andere Sprache) und die aktuellen Schulnoten in allen typischerweise an einer Sekundarschule unterrichteten Schulfächer erfragt wurde.

Abschließend wurden alle Versuchsteilnehmer/-innen vollständig über den Inhalt und das Ziel der Studie aufgeklärt sowie ein symbolischer Scheck mit der versprochenen Entlohnung überreicht. Keine Person äußerte nach mehrmaliger Nachfrage einen ergänzenden Gesprächsbedarf. Insgesamt dauerte die Befragung incl. Pause ca. 90 (± 10) Minuten. Das Vorgehen sowie die mündlich zu gebenden Instruktionen waren mittels eines Untersuchungsleitfadens vollständig standardisiert.

6.3. Ergebnisse

Berechnung der Leistungsdaten und initiale Leistung im Pretest. Für beide Testteile, den Pre- und den Posttest sowie für beide untersuchte Domänen (Deutsch und Mathematik) wurden separat die jeweils korrekt beantworteten Fragen ausgezählt und mit der Gesamtanzahl der Fragen in ein prozentuales Verhältnis gesetzt. Hieraus resultiert für jede Versuchsperson ein Prozentwert je Testteil und Domäne, der den Anteil richtig gelöster Aufgaben wiedergibt. Mit einem Cronbach's $\alpha = 0.66$ für die Aufgaben zum Leseverständnis und $\alpha = 0.67$ für die mathematischen Aufgaben wurde *sensu* Cohen (1988) eine ausreichende Reliabilität erreicht, welche insbesondere vor dem Hintergrund der großen interindividuellen Variabilität von Leistungsdaten für eine reliable Erfassung der Lese- und Mathematikkompetenz spricht.

Um sicherzustellen, dass die Probanden/-innen ein initial vergleichbares Leistungsniveau in beiden Domänen aufwiesen, d.h. dass zwischen den Experimentalgruppen oder Geschlechtern keine bedeutsamen Leistungsunterschiede bestanden, wurden Varianzanalysen (ANOVA) für die Leistung im Pretest in den verschiedenen Domänen in Abhängigkeit von der Gruppenzugehörigkeit und dem Geschlecht berechnet. Zwischen den Experimentalbedingungen lassen sich keine bedeutsamen Leistungsunterschiede in beiden Domänen festhalten, für Lesen: $M_{EG} = 57.47$, $SD = 15.63$, $M_{KG} = 56.11$, $SD = 16.84$; für Mathematik: $M_{EG} = 39.53$, $SD = 18.39$, $M_{KG} = 37.54$, $SD = 17.28$; alle $F < 1$, n.s. Auch zwischen den Geschlechtern lassen sich ebenfalls keine bedeutsamen Leistungsdifferenzen festhalten, Lesen: $M_{Jungen} = 57.94$, $SD = 15.63$; $M_{Mädchen} = 55.69$, $SD = 16.77$; Mathematik: $M_{Jungen} = 39.09$, $SD = 16.15$; $M_{Mädchen} = 38.01$, $SD = 19.36$;

alle $F < 1$, n.s.; zusätzlich trat kein signifikanter Interaktionseffekt Experimentalbedingung \times Geschlecht in Deutsch und Mathematik auf, $F < 1$, n.s.

Somit kann festgehalten werden, dass sich sowohl Experimental- und Kontrollgruppe als auch Jungen und Mädchen nicht hinsichtlich ihrer initialen Leistung in Deutsch und Mathematik unterschieden.

Manipulation Check. Wie bereits beschrieben wurden 20 Probanden und 4 Probandinnen aus den Analysen ausgeschlossen, da diese z.T. bei den vier Kontrollfragen zum Treatment *keine* Frage richtig beantworteten⁵. Die verbleibenden Probanden/-innen ($N = 100$) wiesen mindestens 50% richtige Antworten bei den Fragen zur Manipulationsüberprüfung auf. Im Mittel wurden $M = 72.50\%$ ($SD = 19.94\%$, $range = 50-100\%$) der Kontrollfragen richtig beantwortet, was einem zufriedenstellenden Anteil an korrekt beantworteten Fragen zur Manipulationsüberprüfung entspricht. Eine Varianzanalyse mit den korrekten Prozentanteilen der Kontrollfragen als abhängige Variable und den Faktoren Untersuchungsbedingung und Geschlecht zeigte, dass sowohl zwischen den Untersuchungsbedingungen als auch zwischen den Geschlechtern keine bedeutsamen Haupt- oder Interaktionseffekte hinsichtlich der Kontrollfragen zum Aufgabenverständnis vorlagen (alle $F < 1$, $p > .64$), was dafür spricht, dass sowohl zwischen Experimental- ($M = 73.50$, $SD = 20.46$) und Kontrollgruppe ($M = 71.50$, $SD = 19.57$) als auch zwischen Jungen ($M = 72.96$, $SD = 21.55$) und Mädchen ($M = 72.06$, $SD = 18.47$) keine bedeutsamen Unterschiede bei den Anteilen korrekt beantworteter Kontrollfragen festgestellt werden konnten und in allen Untersuchungsbedingungen eine vergleichbar gute Verarbeitung des Treatment- und Kontrollmaterials erfolgte.

Analyse der Leistungsdaten. Die umfassendste Analyse der Leistungsdaten würde eine $2 \times 2 \times 2$ -faktorielle *mixed model*-ANOVA (Experimentalbedingung \times Geschlecht der Versuchspersonen \times Messzeitpunkt \times Leistungsdomäne) mit Messzeitpunkt (Pre- vs. Posttest) und Leistungsdomäne (Deutsch vs. Mathematik) als Messwiederholungsfaktoren darstellen. Die Ergebnisse zeigten allerdings, dass durch die mentale Belastung der Versuchspersonen in der Testsituation die Leistung generell vom Pre- zum Posttest abfällt und zwar vollkommen unabhängig von irgendeinem anderen Faktor, welcher in die Varianzanalyse aufgenommen wurde. Demzufolge besteht keine Möglichkeit, eine signifikante 4fach-Interaktion nachzuweisen, $F < 1$, n.s. Auch nach der Bildung von Differenzwerten (Leistung Posttest – Leistung Pretest) zur Einsparung eines Messwiederholungsfaktors zeigte sich ein kongruentes Bild, sodass erneut die abfallenden Leistungen unabhängig von allen anderen aufgenommenen Faktoren eine signifi-

⁵Ferner wiesen die ausgeschlossenen Versuchsteilnehmer/-innen weitere Auffälligkeiten auf. Sehr häufig wurden die Testhefte mit Kritzeleien und teilweise abschätzigen Zeichnungen und/oder Kommentaren versehen (z.B. „kein Bock!“). Auch Antwortmuster bei den geschlossenen Fragen waren bei diesem Personenkreis häufig anzutreffen (z.B. ausschließliches Ankreuzen der mittleren Antwortkategorie oder Ankreuzen von Mustern).

kante Dreifachinteraktion mit $F < 1$, n.s. verhinderten. Aus diesem Grund wurden die Analysen für die Leistungsdaten getrennt für die beiden Leistungsdomänen vorgenommen und werden so auch im Folgenden vorgestellt.

Die mittleren Leistungen in den *Leseaufgaben* sind in Abbildung 6 zusammengefasst und wurden einer 2 (Geschlecht der Versuchsperson) \times 2 (Experimentalbedingung) \times 2 (Messzeitpunkt)-faktoriellen *mixed model*-ANOVA mit Messzeitpunkt (Prä- vs. Posttest) als Messwiederholungsfaktor unterzogen. Hier zeigte sich eine statistisch bedeutsame Tripel-Interaktion mit einer mittleren bis großen Effektstärke (Fritz, Morris & Richler, 2012) zwischen allen in die Analyse aufgenommenen Faktoren, $F(1,96) = 10.43$, $p < .01$, $\eta^2 = .10$. Wie vorhergesagt wurde, fällt die Leistung von Jungen in der Experimentalgruppe, also nach dem Lesen des Artikels über Jungen als Schulversager, wesentlich stärker vom Pretest ($M = 57.41\%$) zum Posttest ($M = 27.31\%$) ab, als bei Jungen in der Kontrollbedingung, welche über Veränderungen im deutschen Schulsystem lasen (Pretest: $M = 58.59\%$; Post-test: $M = 40.91\%$). Der Leistungsabfall war ebenfalls deutlicher als bei Mädchen in der Experimentalgruppe (Pretest: $M = 57.53\%$; Posttest: $M = 49.73\%$) oder der Kontrollgruppe (Pretest: $M = 54.17\%$; Posttest: $M = 37.28\%$). Der Leistungsabfall im Posttest betrug in Differenzwerten ausgedrückt: $M_{\text{Jungen_EG}} = -30.09\%$, $M_{\text{Jungen_KG}} = -17.68\%$, $M_{\text{Mädchen_EG}} = -7.81\%$, $M_{\text{Mädchen_KG}} = -16.89\%$.

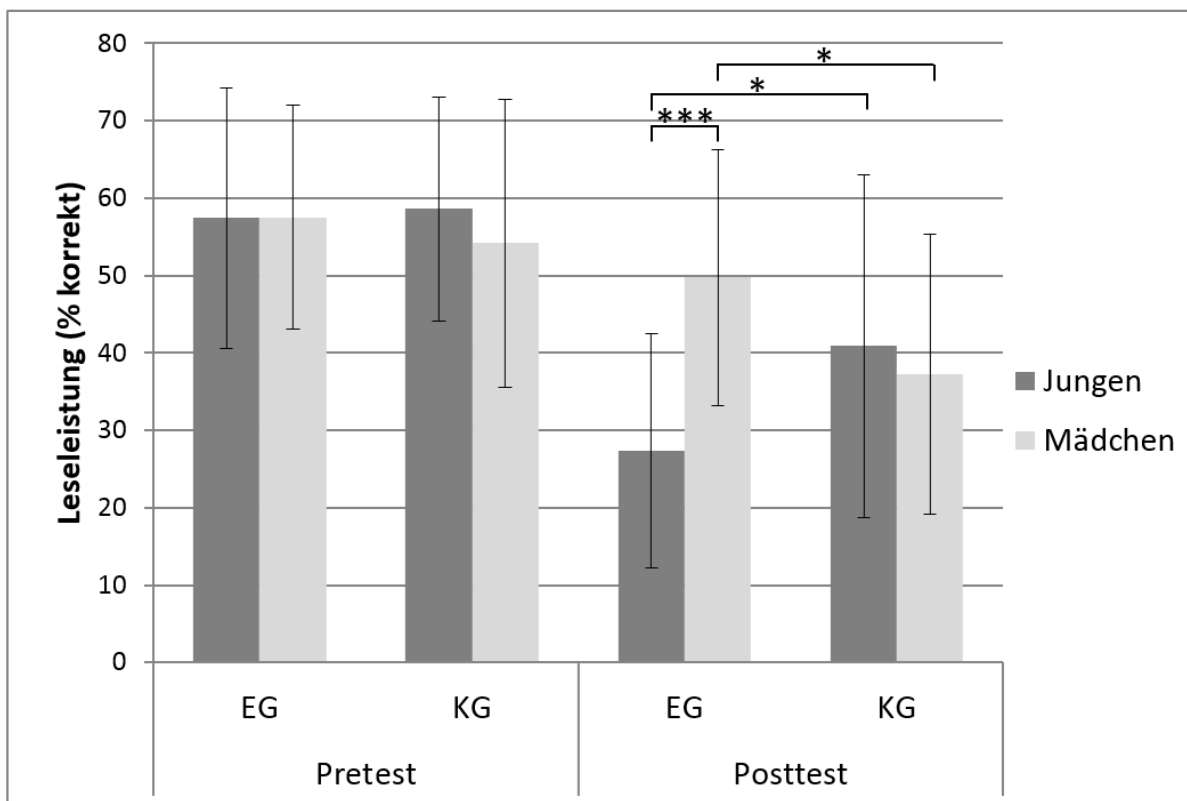


Abbildung 6: Deutsche Leistung von Jungen und Mädchen in der Experimental- und Kontrollbedingung im Prä- und Posttest. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Anm.: genereller Leistungsabfall in *allen* Untersuchungsgruppen vom Pre- zum Posttest (ohne Kennzeichnung des Signifikanzniveaus).

Post-hoc Einzelvergleiche mittels T-Tests⁶ verdeutlichten, dass im Pretest keine bedeutsamen Unterschiede bestanden, sich im Posttest jedoch Jungen und Mädchen der Experimentalgruppe bedeutsam in ihrer Leseleistung – im Sinne einer deutlich besseren Leseleistung von Mädchen – unterschieden, $t(48) = 5.01, p < .001, d = -1.42$. Ebenfalls erreichten im Posttest Jungen in der Kontrollgruppe bedeutsam bessere Werte hinsichtlich der Leseleistung, als Jungen in der Experimentalbedingung, $t(47) = -2.45, p < .05, d = -0.72$. Ergänzend zeigte sich ein bedeutsamer Effekt für Mädchen in der Experimentalbedingung im Vergleich zu Mädchen der Kontrollbedingung, $t(49) = 2.54, p < .05, d = 0.72$; d.h. es lagen bei Mädchen dann bessere Leseleistungen vor, wenn sie zuvor den Artikel über die bildungsschwachen Jungen gelesen haben⁷.

Die Dreifachinteraktion wurde von einem signifikanten Haupteffekt des Messzeitpunktes begleitet, der deutlich machte, dass unabhängig von Geschlecht oder experimenteller Bedingung die Leistung im Posttest mit einer substanziell großen Effektstärke schlechter ausfiel, $M = 38.25\%$, als im Pretest $M = 56.79\%$, $F(1,96) = 118.50, p < .001, \eta^2 = .55$ (vgl. auch Abbildung 6). Ferner wurde eine bedeutsame Interaktion zwischen den Variablen Messzeitpunkt und Geschlecht gefunden, $F(1,92) = 12.01, p = .001, \eta^2 = .11$. Während Jungen, $M = 57.94\%$, und Mädchen, $M = 55.69$, im Pretest vergleichbare Ergebnisse im Lesen erzielten, war die Leistung der Jungen im Posttest sehr viel geringer, $M = 33.41$, als die durch Mädchen erzielten Ergebnisse, $M = 42.90$. Zu berücksichtigen bleibt bei diesem Effekt allerdings, dass dieser sich in der Hauptsache über die stark nachlassende Leistung der Jungen in der Experimentalbedingung qualifizierte, weshalb auf eine genauere Interpretation verzichtet wurde.

Für den *mathematischen Test* wurde kongruent zur Auswertung der Deutschaufgaben eine 2 (Geschlecht der Versuchsperson) $\times 2$ (Experimentalbedingung) $\times 2$ (Messzeitpunkt)-faktorielle *mixed model*-ANOVA mit Messzeitpunkt (Prä- vs. Posttest) als Messwiederholungsfaktor berechnet. Die Verteilung der Mittelwerte nach Testphase, Experimentalbedingung und Geschlecht können der folgenden Abbildung entnommen werden.

⁶Für andere post hoc-Verfahren wie den Scheffé-Test oder Tukey-Test gilt die Voraussetzung, dass mindestens drei Gruppen für die Berechnung der post-hoc-Statistiken vorhanden sein müssen. Da die in die o.g. ANOVA aufgenommenen Variablen allerdings nur zweifach gestuft sind, bieten sich diese Verfahren nicht an, weshalb auf den T-Test zurückgegriffen wurde.

⁷ Werden die post-hoc T-Tests nach Bonferoni korrigiert entspricht das Fehlerniveau bei vier möglichen Tests einem $\alpha = .0125$. Wird dieses Kriterium angelegt, erreicht lediglich der Einzelvergleich im Posttest von Jungen und Mädchen in der Experimentalgruppe statistische Bedeutsamkeit ($p < .001$), die anderen Einzelvergleiche fallen mit $p = .019$ und $p = .014$ knapp über die adjustierte Fehlerwahrscheinlichkeit.

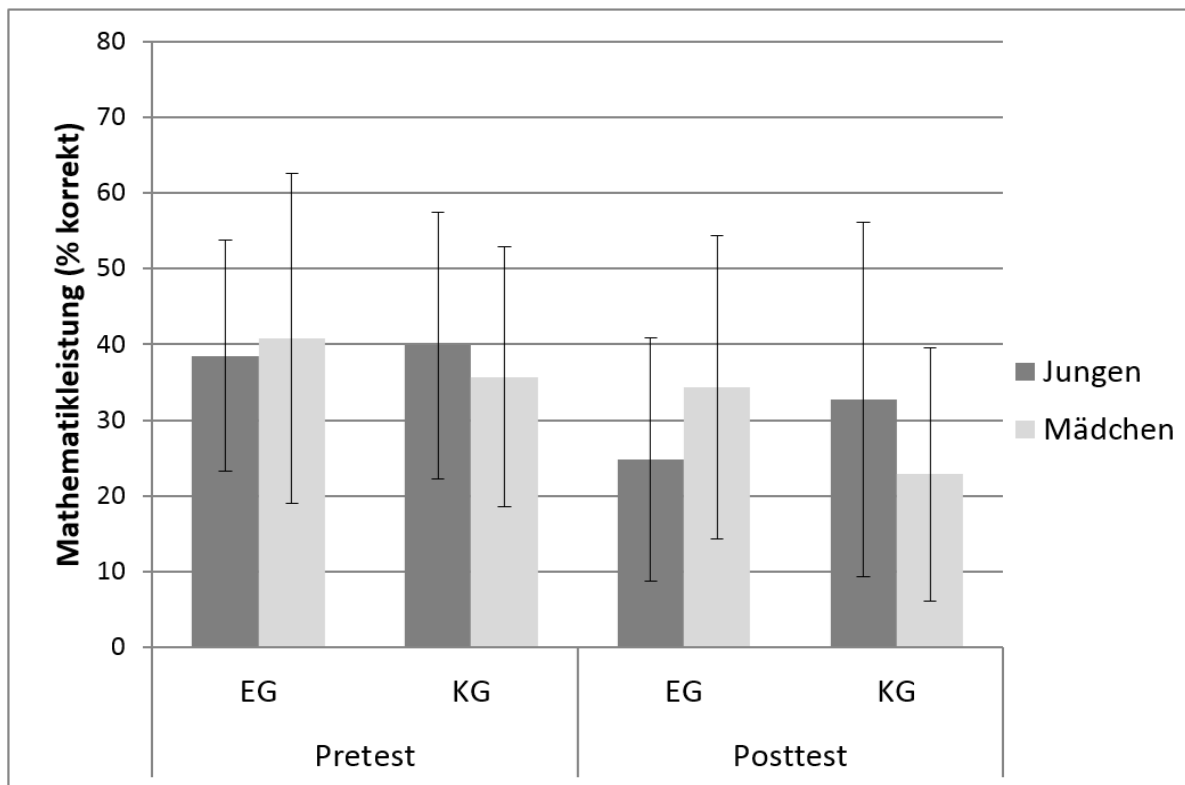


Abbildung 7: Mathematikleistung von Jungen und Mädchen in der Experimental- und Kontrollbedingung im Prä- und Posttest. * = $p < 0.05$, ** = $p < 0.01$ *** = $p < 0.001$

Anm.: genereller Leistungsabfall in *allen* Untersuchungsgruppen vom Pre- zum Posttest (ohne Kennzeichnung des Signifikanzniveaus).

Deskriptiv fällt auf, dass das Leistungsniveau bei den Mathematikaufgaben im Vergleich zu den Leseaufgaben geringer ausfiel. Die Analyseergebnisse der ANOVA zeigten einen signifikanten Haupteffekt für die Variable Messzeitpunkt, $F(1,96) = 22.54$, $p < .001$, $\eta^2 = .19$, was auf eine erneut schlechtere Leistung im Posttest ($M = 28.20\%$) im Vergleich zum Pretest ($M = 38.54\%$) hindeutete. Weder der Interaktionsterm zwischen Messzeitpunkt und Geschlecht, noch die Interaktion zwischen Messzeitpunkt und Experimentalbedingung erreichten eine statistische Bedeutsamkeit, beide $F < 1$, n.s. Auch die Tripel-Interaktion erreichte kein substanzielles Signifikanzniveau, $F(1,96) = 2.34$, $p = .13$. Insgesamt kann somit festgehalten werden, dass außer einem generellen Abfall der Leistung im Posttest unabhängig vom Geschlecht der Versuchsperson oder der Untersuchungsbedingung keinerlei bedeutsame Unterschiede hinsichtlich der Mathematikleistung festgestellt werden konnten. Die Differenzwerte (Posttest – Pretest, negative Werte bedeuten somit einen Leistungsabfall) betragen: $M_{Jungen_EG} = -13.65\%$, $M_{Jungen_KG} = -7.13\%$, $M_{Mädchen_EG} = -6.45\%$, $M_{Mädchen_KG} = -12.86\%$. Aufgrund der ausbleibenden signifikanten Haupt- und Interaktionseffekte wurde auf post hoc-Testungen verzichtet.

Um auf die Abhängigkeit der Lösungswahrscheinlichkeit einer Aufgabe von der Aufgabenschwierigkeit und individueller Kompetenz angemessen reagieren zu können, wurden

die Leistungsdaten basierend auf der *Item-Response-Theorie* zusätzlich einer dichotomen Raschmodellierung unterzogen (Kempf, 1974; Rasch, 1960; Rost, 2003). Die Partialisierung der verwendeten Items zur Leistungsmessung in einen Personenparameter θ_v und einen Aufgabenparameter σ_i führten zur Selektion von insgesamt jeweils 10 Items für jede Leistungsdomäne sowie die Prä- und Posttestung, für welche das getestete Rasch-Modell seine Gültigkeit beibehielt, Deutsch: $\chi^2 = 17.38$, $p = .56$, $Wald\ z = |0.01| - |1.59|$, $p > .11$; Mathematik: $\chi^2 = 16.80$, $p = .60$, $Wald\ z = |1.80| - |0.01|$, $p > .07$. Aufbauend auf diesen Analysen konnte gezeigt werden, dass mit den ausgewählten Items ein Vergleich der Probanden/-innen unabhängig von den Aufgabencharakteristika und die methodisch unproblematische Bildung von Differenzwerten ($P_{\text{Post-Test}} - P_{\text{Pre-Test}}$) zur Analyse der Leistungsentwicklung vom Pre- zum Posttest möglich wurde. Weiterhin konnte ein grafischer Test der Itemcharakteristikkurven (*ICC's*) deutlich machen, dass die Aufgaben des Pre- und Posttests hinsichtlich der Lösungswahrscheinlichkeit vergleichbar ausfielen, sodass von zwei hochgradig vergleichbaren Leistungsmessungen in beiden untersuchten Domänen ausgegangen werden konnte. Weiterführende Analysen mit den durch die Raschmodellierung selektierten Items der Leistungsmessung bestätigten die o.g. Befunde im Rahmen von 2 (Geschlecht) \times 2 (Untersuchungsbedingung) – Varianzanalysen zum Einfluss des experimentellen Treatments auf die Leistung von Jungen vollumfänglich: In beiden untersuchten Domänen unterschieden sich die Leistungen im Pretest weiterhin nicht bedeutsam voneinander, weder zwischen den Untersuchungsgruppen im Lesen, $M_{EG} = 56.80\%$, $SD = 17.89\%$; $M_{KG} = 53.80\%$, $SD = 18.83\%$, noch in Mathematik: $M_{EG} = 46.00\%$, $SD = 20.90\%$, $M_{KG} = 42.40\%$, $SD = 18.69$; alle $F < 1$, n.s. Ebenso gab es keine bedeutsamen Unterschiede zwischen den Geschlechtern im Pretest in beiden Leistungsbereichen; Lesen: $M_{\text{Jungen}} = 55.71\%$, $SD = 17.91\%$; $M_{\text{Mädchen}} = 54.90\%$, $SD = 18.91\%$; Mathematik: $M_{\text{Jungen}} = 43.27\%$, $SD = 19.30\%$; $M_{\text{Mädchen}} = 45.10\%$, $SD = 20.43\%$; alle $F < 1$, n.s.

Im *Lernbereich Deutsch* deutete ein signifikanter Haupteffekt, $F(1,96) = 4.84$, $p < .05$, $\eta^2 = .05$, für den Faktor Geschlecht darauf hin, dass Jungen einen stärkeren Leistungseinbruch ($M = -20.00\%$) im Lese-Posttest aufweisen als Mädchen ($M = -10.39\%$). Dieser Haupteffekt qualifizierte sich allerdings über eine signifikante Interaktion aus den Faktoren Geschlecht und Untersuchungsbedingung, $F(1,96) = 7.92$, $p < .01$, $\eta^2 = .08$. Wie in den vorausgegangenen Analysen mit den nicht-raschskalierten Daten zeigten Jungen der Experimentalgruppe im Lesen nach der Bearbeitung des *stereotype threat*-Treatments eine bedeutsam verringerte Leseleistung ($M = -27.04\%$) im Vergleich zu Jungen der Kontrollgruppe ($M = -11.36\%$) oder Mädchen der Experimental- ($M = -6.09\%$) und Kontrollgruppe ($M = -13.93\%$), post-hoc Einzelvergleiche: Jungen Experimental- vs. Kontrollgruppe: $t(47) = -2.79$, $p < .01$, $d = -.81$; Jungen vs. Mädchen in der Experimentalgruppe: $t(41) = -3.57$, $p = .001$, $d = -1.01$.

Im *Lernbereich Mathematik* zeigten sich – wie in den vorausgegangenen Analysen mit den Rohdaten – keinerlei bedeutsame Unterschiede zwischen den Geschlechtern (M_{Jungen}

= -16.33%, $M_{\text{Mädchen}} = -19.02\%$, $F < 1$, n.s.), den Untersuchungsbedingungen ($M_{\text{EG}} = -17.60\%$, $M_{\text{KG}} = -17.80\%$, $F < 1$, n.s.), oder bezüglich der Interaktionseffekte aus beiden Faktoren, $F(1,96) = 1.56$, $p < 0.22$. Somit konnte gezeigt werden, dass auch unter verstärkter statistischer Kontrolle von möglichen Einflüssen der Aufgabencharakteristika des Leistungstests die initial gefundenen Auswirkungen einer Stereotypenbedrohung für Jungen nur im Fach Deutsch aufrechterhalten blieben und in der Mathematik nicht nachzuweisen waren.

6.4. Diskussion

Die Ergebnisse dieser Studie zeigen, dass Jungen, welche einen Artikel über das schlechtere schulische Abschneiden der Eigengruppe gelesen haben nur im Fach Deutsch, nicht jedoch im Fach Mathematik leistungsmäßig stärker nachlassen als Mädchen oder aber Jungen der Kontrollbedingung, in welcher ein geschlechtsneutraler, unbedrohlicher Lesetext zu bearbeiten war. Im Fach Deutsch konnte zusätzlich festgestellt werden, dass Mädchen von der Darstellung von Jungen als Bildungsverlierer in ihrer Leistung sogar profitieren konnten und eine vergleichsweise bessere Leistung zeigten, als Probandinnen der Kontrollbedingung. Zusätzlich lässt sich dieses Bild zumindest auf deskriptiver Ebene (aber leider ohne statistischen Effekt) auf das Fach Mathematik übertragen: Auch hier weisen Mädchen nach dem Lesen des die Außengruppe bedrohenden Artikels deskriptiv bessere Leistungen auf als Mädchen in der Kontrollgruppe. Zusammenfassend kann also festgehalten werden, dass durch die Applikation einer generellen, domänenübergreifenden Stereotypenbedrohung die Leistung von Jungen keineswegs ebenso generell beeinträchtigt wurde, sondern sich die leistungshemmenden Effekte nur im Fach Deutsch niederschlugen. Der Profit von Mädchen aus der negativen Darstellung von Jungen in der Schule erreichte ebenso nur im Fach Deutsch eine statistische Bedeutsamkeit.

Bezugnehmend auf die Hypothesen lässt sich somit festhalten, dass Hypothese 4 zumindest teilweise bestätigt werden konnte, da ein domänenspezifischer Effekt für das Fach Deutsch in vorhergesagter Richtung festgestellt wurde, dies aber im Kontrast nicht für Mathematik gilt. Für die Hypothese 5 existiert Evidenz, da die Ergebnisse für die Vorhersage sprechen, dass Mädchen als Mitglieder der nicht-stereotypisierten Außengruppe von der abwertenden Darstellung bezüglich der schulischen Leistungsfähigkeit von Jungen profitieren. Eine umfassende Bestätigung wie vorhergesagt blieb allerdings kongruent zur vorhergehenden Hypothese aus, da nur im Fach Deutsch eine statistisch abgesicherte, bessere Leistung von Mädchen der Experimentalbedingung (im Vergleich zu Mädchen der Kontrollbedingung) aufgezeigt werden konnte. Weiterhin konnte festgestellt werden, dass die Leistung von Jungen und Mädchen der Kontrollbedingung sich in beiden untersuchten Domänen weder im Prä-, noch im Posttest unterschied. Neben einem (auch in den anderen Untersuchungsgruppen)

festgestellten Leistungsabfall vom Prä- zum Posttest divergieren die Leistungen innerhalb der Kontrollgruppe nicht bedeutsam zwischen den Geschlechtern.

Somit ist festzuhalten, dass die Auswirkungen der generellen Stereotypenbedrohung anders ausfielen, als dies im Rahmen der Hypothesen vorhergesagt wurde. Statt generellen, domänenübergreifenden Resultaten konzentrieren sich die Auswirkungen auf Ergebnismuster, welche die generell vorherrschenden Geschlechterstereotype in den untersuchten Schulfächern widerspiegeln (Plante et al., 2009). Es kann diskutiert werden, inwieweit der Glaube an das Stereotyp hier zum Tragen kommt (vgl. Schmader et al., 2004), d.h. inwiefern sich Jungen tatsächlich übergreifend, in allen schulischen Fächern als unterlegen wahrnehmen. Ferner kann diskutiert werden, warum bei bedrohten Jungen im Fach Mathematik die generelle Bedrohung keine Auswirkungen hat. Die umfassende Bedrohung von Jungen als Schulversager *across the board* spiegelt u.U. nicht ausreichend das wieder, was Jungen im Schulalltag selbst erleben und wahrnehmen. Möglicherweise liefern leistungsbezogene Geschlechterstereotype hier ein genaueres Bild wieder und führen zu den berichteten Ergebnissen sowie dazu, dass Jungen sich in „ihrer“ Domäne Mathematik gegen die negative Verhaltenserwartung durch die Darstellung als generelle Minderleister zur Wehr gesetzt haben (Kray et al., 2001, 2004). Was in den Medien kommuniziert wird glauben die Schüler/-innen zwar inzwischen in ebenso generellem Ausmaß, wie es dargestellt wird auch selbst (siehe *Studie 1*), allerdings kann die Auswirkung dieser unspezifischen Bedrohungssituation in der vorliegenden Studie offenkundig vielmehr als Mechanismus zur Aufrechterhaltung der bisher geltenden leistungsbezogenen Geschlechterstereotype angesehen werden.

Hinsichtlich der Einordnung der vorliegenden Befunde in den existierenden, empirischen Forschungsstand zeigt sich ein gemischtes Bild. Es existiert einerseits Evidenz für einen leistungshemmenden Effekt von *stereotype threat* in einem generellen, domänenübergreifenden Ausmaß konträr zu den domänenspezifischen Effekten in der vorliegenden Studie (Hartley & Sutton, 2013). In dem Altersspektrum, in welchem die Untersuchung von Hartley & Sutton (2013) angesiedelt wurde, scheint es demnach so zu sein, dass die applizierte, generelle Stereotypenbedrohung wesentlich rigider angewendet wurde (Trautner, Ruble, Cyphers, Kirsten, Behrendt & Hartmann, 2005), als es in der vorliegenden Stichprobe von Jugendlichen der Fall zu sein scheint. Mit der fortgeschrittenen Entwicklung und der damit verbundenen, steigenden Erfahrung im Lernkontext Schule scheint es mit einem fortgeschrittenen Alter zu einer Ausdifferenzierung der generellen Stereotypenbedrohung gekommen zu sein, in dem Sinne, dass ein Leistungsabfall nur in der Domäne hervorgebracht werden konnte, in welchem die gängigen Geschlechterstereotype gegen den Erfolg der *Ingroup* spricht. In anderen Bereichen, in welchen die generelle Bedrohung als illegitim wahrgenommen wird, da hier die

Geschlechterstereotypen eigentlich für die Leistung der Eigengruppe spricht, scheint keine Vulnerabilität gegenüber *stereotype threat* vorzuliegen oder aber eine Reaktanzreaktion vorherrschend zu sein. Weiterhin wurde gezeigt, dass ein induzierter *prevention focus* mit einer verminderten Testleistung, ein induzierter *promotion focus* mit einer verbesserten Testleistung unter Stereotypenbedrohung einhergeht (Keller, 2007). Die Ergebnisse der Studie 2 finden ohne Trennung nach regulatorischen Foki einen leistungshemmenden Effekt von *stereotype threat* und stimmen dahingehend mit denen der vorliegenden Studie überein, als dass gezeigt werden konnte, dass ohne bewusste Manipulation des regulatorischen Fokus eine Stereotypenbedrohung spontan eher mit einem *prevention focus* einhergeht (Seibt & Förster, 2004), aus welchem die verminderte Leistung, wie sie auch in der vorliegenden Studie gezeigt wurde, gespeist wird.

Konträr zu diesen Befunden wird argumentiert, dass der schädliche Effekt einer Stereotypenbedrohung, welcher bei Frauen in mathematisch-naturwissenschaftlichen Bereichen als nachgewiesen gilt (Ambady, Shih, Kim & Pittinsky, 2001; Keller, 2002; Nguyen & Ryan, 2008) nicht *per se* auf Jungen im sprachlichen Bereich übertragen werden kann, da keinerlei leistungsmindernde Effekte durch eine Bedrohung nachgewiesen werden konnten (Eckert & Imhof, 2013). Aufgrund zahlreicher, methodischer Probleme⁸ bleibt allerdings zu hinterfragen, ob die fehlenden, leistungshemmenden Wirkungen der Stereotypenbedrohung tatsächlich als Evidenz für die Insensitivität von Jungen gewertet werden dürfen.

Zusätzlich existieren Studien, in welchen sogar ein leistungssteigernder Effekt durch die Stereotypenbedrohung innerhalb der bedrohten Gruppe gefunden werden konnte, was den vorliegenden Ergebnissen widerspricht. Hirnstein et al. (2012) finden in einem Wortflüssigkeitstest eine gesteigerte Leistung von Männern unter *stereotype threat* und interpretieren dies im Sinne einer Wettbewerbssituation oder als Vorliegen von *stereotype reactance*. Hier ist relevant, dass sich die Untersuchung von Hirnstein et al. (2012) auf nur einen Fähigkeitsbereich bezieht, während die vorliegende Untersuchung mehrere Fähigkeitsbereiche untersucht, für welche die vorherrschenden Geschlechterstereotype sowohl für als auch gegen die Leistungsfähigkeit von Jungen spricht. Im Sinne einer postulierten Reaktanz könnte dies die

⁸ Zwei Drittel der befragten Personen aus Eckert & Imhof (2013) entstammen einer gymnasialen Stichprobe. Es ist fraglich, inwiefern diese Probanden vulnerabel auf die Stereotypenbedrohung im dem Bewusstsein reagieren, die höchste deutsche Schulform zu besuchen, was zusätzlich mit einem substanziellen Kompetenzvorsprung einhergeht: In Deutsch werden 77 Kompetenzpunkte mehr als in der Realschule erreicht, in Mathematik sogar 83 Kompetenzpunkte (Klieme et al., 2010). Beides entspricht dem Kompetenzunterschied von ca. 2 Schuljahren (OECD, 2014). Inwiefern daher die negativen Effekte von *stereotype threat* durch diesen substanziellen Anteil an der Gesamtstichprobe konterkariert werden oder aber *stereotype reactance* als Konsequenz der Bedrohung ausgelöst wird, geht nicht aus der Studie hervor. Da Effekte durch *stereotype threat* vornehmlich bei schwierigen Aufgaben auftreten (u.a. Spencer et al., 1999; O'Brien & Crandall, 2003; Neuville & Croizet, 2007) kann weiterhin diskutiert werden inwiefern die Aufgaben für die untersuchten Gymnasiasten hinreichend schwierig zu lösen waren. Ferner werden nur Jungen analysiert, wengleich die Kontrastierung beider Geschlechter eine wünschenswerte Bedingung darstellt (Stoet & Geary, 2012).

Möglichkeit eröffnen, dass sich diese durch die Einführung einer weiteren, positiv stereotypisierten Domäne entsprechend verlagert und nicht in leistungssteigernde Effekte im negativ stereotypisierten Fachbereich führt. Dagegen spricht allerdings, dass somit eine deutliche Leistungssteigerung in der positiv stereotypisierten Domäne erwartbar gewesen wäre, welche in der vorliegenden Untersuchung allerdings nicht nachgewiesen werden konnte.

Insgesamt kann zusammenfassend festgehalten werden, dass sich unter Bezugnahme der wenigen, verfügbaren Studien zu Jungen und *stereotype threat* andeutet, dass die gleiche schädliche Wirkung einer Stereotypenbedrohung, wie sie für Frauen im MINT-Bereich nachgewiesen wurde, auch bei Jungen im Fach Deutsch, also in der Domäne, in welcher das Geschlechterstereotyp gegen die Leistungsfähigkeit von Jungen spricht, auftreten kann. Zusätzlich wurde eine generellere Bedrohung eingesetzt, als dies bisher der Fall war. Hierbei zeigt sich, dass diese im Kindesalter eine ebenso globale Reaktion auslösen (Hartley & Sutton, 2013) und sich im Jugendalter im Sinne einer Reduzierung der generellen Bedrohung anhand der vorherrschenden Geschlechterstereotypen ausdifferenzieren und bei wahrgenommener Illegitimität der Bedrohung Reaktanz (Kray et al., 2001) nach sich ziehen kann. Somit kann vermutet werden, dass der Bedrohungsreiz im Jugendalter u.U. nicht zu allgemein sein darf, da sonst andere Verarbeitungsmechanismen angestoßen werden, wie hier z.B. die Aktivierung von Geschlechterstereotypen. Für diese Interpretation spricht ferner, dass sich die fächerspezifischen Selbstkonzepte von Schüler/-innen im Verlauf der Schulzeit mehr und mehr ausdifferenzieren (Shavelson, Hubner & Stanton, 1976; Marsh, 1990; Marsh & Ayotte, 2003). Während in der frühen Schulzeit also noch eine Bedrohung mit „schlechteren Schulleistungen“ erreicht werden kann, ist dies im Jugendalter mit einem vergleichsweise sehr ausdifferenzierten, fachspezifischen Selbstkonzept deutlich schwerer, eine generelle Bedrohung zu applizieren, weil schulische Leistungen vielmehr als Teilbereiche und weniger als globales Konstrukt angesehen werden. Zusätzlich erfolgt während der Entwicklung vom Kindes- zum Jugendalter eine verstärkte Ausformung der Inhalte von Geschlechterstereotypen (Eckes, 2008; Zemore, Fiske & Kim, 2000), was derartige Effekte weiter unterstützen kann.

Die Ergebnisse der vorliegenden Studie ergänzen somit die bisherigen Forschungsergebnisse auf folgende Art und Weise: Es wurde nicht nur ein Stereotyp, sondern eine breitere Bedrohung im Sinne eines globalen *stereotype threat* appliziert. Zusätzlich wurden nicht nur Aufgaben in einer für Jungen negativ konnotierten Leistungsdomäne, sondern auch Aufgaben mit einer positiven Leistungserwartung entsprechend der geltenden, leistungsbezogenen Geschlechterstereotype zur Bearbeitung vorgelegt. Somit konnte ergänzend zu bisherigen Befunden gezeigt werden, dass eine generelle Stereotypenbedrohung von Jungen im Jugendalter, wie die aktuelle mediale Berichterstattung sie vornimmt, nicht im Sinne einer generell negativen

self-fulfilling prophecy (Merton, 1948) funktioniert, sondern vielmehr eine differenziertere Wirkungsweise entfaltet.

Limitationen. Auffällig bei der beschriebenen Untersuchung ist der vergleichsweise hohe Anteil von Jungen, welche aus den weiterführenden Analysen ausgeschlossen werden mussten – welche sich allerdings gleich auf beide Untersuchungsbedingungen aufteilen. Es bleibt im Rahmen dieser Untersuchung daher zunächst ungeklärt, was genau zu diesem vergleichsweise hohen Ausschluss von Jungen geführt hat, welche offensichtlich die Untersuchung durch das Ankreuzen von Mustern, eine unsorgfältige Bearbeitung der Leistungsaufgaben und das zahlreiche Auslassen von Fragen boykottiert haben. Es kann nur vermutet werden, dass die generelle Distanz zur Schule (Segeritz et al., 2010) und den damit verbundenen Aufgaben (Spiel et al., 2002) zu einer generell reduzierten Bereitschaft geführt haben, an der Testsituation aktiv teilzunehmen. Durch den fehlenden Einfluss der Untersuchungsbedingung kann nicht geschlussfolgert werden, dass die Ausführungen im Treatment („schlaue Mädchen, dumme Jungen“) zu einem Effekt im Sinne einer Leistungsverweigerung geführt haben. Auch kann eine mögliche Überforderung mit den Inhalten des Leistungstests nicht ausgeschlossen werden, die zu einem Abbruch der Bearbeitung und der Herausbildung einer Frustration geführt haben kann (Steele & Aronson, 1995), welche sich wiederum in einer stark verminderten Anstrengung niederschlägt (Bandura, 1986) oder gar zu einer Entladung im Sinne der aufgefundenen, abwertenden Kommentare und Abbildungen führt. Die Erfassung der Einstellungen gegenüber der Testsituation, der Bereitschaft zur gewissenhaften Teilnahme sowie des Erlebens der Testsituation im Rahmen einer Nachbefragung könnten in Zukunft genauere Aussagen über die Beweggründe von Leistungsverweigerern innerhalb der Testsituation erlauben.

Weiterhin wird stark darüber diskutiert, inwiefern ein Einbezug aktueller Leistung, d.h. bestehender Leistungsunterschiede vor der Bedrohungssituation in die Datenanalyse notwendig ist. Für die vorliegende Studie kann festgehalten werden, dass eine ANCOVA zwar auf einen bedeutsamen Einfluss der aktuellen Schulnoten in den untersuchten Fächern hinweist, jedoch der ohne die Kovariate gefundene Effekt in Richtung und Effektgröße kongruent zu den berichteten Ergebnissen ohne Einbezug der Kovariate festgestellt werden konnte. Die Kontrolle dieser bestehenden Leistungsunterschiede wird jedoch teilweise kritisch diskutiert, da diese Analysemethode den eigentlichen Erklärungsversuch dieser initialen Leistungsunterschiede mit der Theorie der Stereotypenbedrohung massiv verletze (Stoet & Geary, 2012). Daher wurde versucht durch das experimentelle Setting (Pre-Post-Test) das initiale Leistungsniveau besser zu kontrollieren.

Der Mathematiktest erreichte gemessen an den korrekten Antworten vergleichsweise schlechtere Werte, die darauf hindeuten, dass die mathematischen Aufgaben a) schwieriger zu beantworten waren *oder* b) in diesem Leistungsbereich ein generell niedrigeres Leistungsniveau

vorherrscht. Es liegen allerdings (noch) keine Bodeneffekte (Bortz & Döring, 2006) vor, sodass eine Interpretation möglich ist. Die Aufgaben wurden sorgfältig entsprechend curricularer Vorgaben gewählt, die Aufgabenschwierigkeiten entsprachen einem durchschnittlichen Anforderungsniveau. Eine noch sorgfältigere Auswahl der Aufgaben, ggf. auch der Einbezug vermehrt einfacher Aufgaben könnte zu einem besser interpretierbaren Datenmuster führen, wenngleich hierbei unbedingt berücksichtigt werden muss, dass Effekte durch *stereotype threat* bei einfachen Aufgaben unwahrscheinlicher werden (Spencer et al., 1999). Abschließend bleibt ungeklärt, wie lange die Wirkung der Stereotypenbedrohung andauert und ob im Anwendungsfeld Schule entsprechend häufig Stereotypenbedrohungen auftreten, sodass die gefundenen Effekte tatsächlich in Zusammenhang mit den bestehenden, überdauernden Leistungsunterschieden gebracht werden können. Ebenso unklar bleibt, inwiefern sich die applizierte Stereotypenbedrohung auf nicht-kognitive, aber dennoch hochgradig lernerfolgsrelevante Faktoren, wie z.B. die motivationale Zielorientierung (Smith, 2004, 2006) auswirkt, was im Rahmen der folgenden Studien genauer betrachtet werden soll.

7. Studien 3 & 4: Stereotype Threat und motivationale Zielorientierungen Auswirkung der medialen Berichterstattung auf einen nicht-kognitiven Faktor

Veröffentlicht als Studie 3 und Replikation in:

Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!?! How the public discourse on failing boys impacts performance outcomes and motivational goal orientation in German students. *Social Psychology, 45*, 112-126.

7.1. Studie 3: Überblick und Fragestellung

Es konnte gezeigt werden, dass Jungen die öffentliche Diskussion über ihr schulisches Scheitern wahrnehmen und internalisiert haben (Studie 1) und daraus eine Bedrohung im Sinne von *stereotype threat* erwachsen kann, welche die Schulleistung spezifisch in den Unterrichtsfächern hemmt, in welchen die leistungsbezogenen Geschlechterstereotypen gegen den Erfolg von Jungen sprechen (Studie 2). Im Folgenden soll überprüft werden, welche Auswirkungen durch eine Stereotypenbedrohung dazu beitragen können, die gefundenen Effekte dauerhaft zu verfestigen und eine kurzfristige Bedrohung in ein überdauerndes Befundmuster zu überführen. Wenn versucht werden soll, mit dem Paradigma der Stereotypenbedrohung auch die überdauernden Leistungsunterschiede zwischen Jungen und Mädchen in der Schule (OECD, 2014; Statistisches Bundesamt, 2014a) zu erklären, muss entweder davon ausgegangen werden, dass Geschlecht und die daraus erwachsende Bedrohung in schulischen Lern- und Leistungskontexten permanent salient sind und so zu einer andauernden Beeinträchtigung der Leistung führen, oder aber, dass durch *stereotype threat* auch nicht-kognitive Bereiche jenseits der aktuellen Leistung beeinflusst werden. Jessie Smith legt mit ihrem *stereotyped task engagement process* (STEP; Smith, 2004) ein Modell vor, welches zur Erklärung von längerfristigen Konsequenzen einer Stereotypenbedrohung beitragen kann und welches in verschiedenen, empirischen Studien validiert und angepasst wurde (u.a. Smith, 2004, 2006; Smith et al., 2007). Zentrale Komponenten dieses Modells sind motivationale Zielorientierungen, welche durch ihr Zusammenspiel mit einer Stereotypenbedrohung die Anwendung vermeidender Lern- und Leistungsstrategien wahrscheinlicher werden lassen. Die Vorhersage des Modells wäre, dass durch eine Stereotypenbedrohung verstärkt Vermeidungs-Leistungsziele einhergehen, welche sich wiederum nachteilig auf die aktuelle Leistung (Smith, 2004), aber auch auf den überdauernden Kompetenzerwerb (Dweck & Leggett, 1988; Elliot, 1999) auswirken. Die Annahme besteht also darin, dass die Auswirkungen einer Stereotypenbedrohung auf Zielorientierungen zur Erklärung beitragen kann, dass trotz der kurzfristigen Expositionsdauer eines bedrohlichen Reizes Effekte auftreten können, die zeitlich überdauernd sind und zu einer Verfestigung führen.

Nach bisherigem Forschungsstand erfolgte bisher noch keine Übertragung des beschriebenen Wirkungsmechanismus auf den Kontext von Jungen in der Schule. Zentrale Fragestellung der vorliegenden Studie 3 soll daher sein, ob die bisherigen Befunde zu *stereotype threat* und Zielorientierungen auf den Kontext von Jungen in der Schule übertragbar sind, um überprüfen zu können, inwiefern die Bedrohung durch die Darstellung von Jungen als Bildungsverlierer Auswirkungen auf einen non-kognitiven Bereich hat und diese ggf. dazu führen können, dass aus Bedrohungsreizen auch langfristige Nachteile im Lernverhalten erwachsen. Unter Bezugnahme auf die Kritik an Eckert & Imhof (2013) ist besonders interessant, diese Untersuchungsfrage auf Jungen in der höchsten deutschen Schulform anzuwenden, um festzustellen, inwiefern hier auf

eine Stereotypenbedrohung mit motivationalen Einbußen oder aber mit Reaktanz reagiert wird. Im Hinblick auf die domänenspezifischen Effekte, die in der vorhergehenden Studie festgestellt wurden, sollen folgende Hypothesen spezifiziert und überprüft werden:

- H6* Jungen in der Experimentalbedingung weisen nach der Interpretation des stereotypenbedrohenden Materials in dem Fach, in welchem das Leistungsstereotyp *gegen* die Leistungsfähigkeit der *Ingroup* spricht, verstärkt eine lernhinderliche Vermeidungs-Leistungszielorientierung im Vergleich zur Kontrollgruppe auf.
- H7* Jungen in der Experimentalbedingung weisen in dem Fach, in welchem das Leistungsstereotyp *für* die Leistungsfähigkeit der *Ingroup* spricht, verstärkt eine lernförderliche Lernzielorientierung im Vergleich zur Kontrollgruppe auf.
- H8* Mädchen in der Experimentalbedingung profitieren von der Darstellung der schulversagenden Jungen nur im Fach Deutsch, in welchem die gängigen Leistungsstereotype *für* Mädchen sprechen und weisen hier eine verstärkte Lernzielorientierung im Vergleich zu Mädchen in der Kontrollbedingung auf.

Die Formulierung gerichteter, domänenspezifischer Hypothesen basiert auf den Befunden aus Studie 2, in welcher festgestellt werden konnte, dass einerseits Jungen durch die Applikation einer generellen Bedrohung nur eine domänenspezifische Leistungseinschränkung aufweisen, welche den vorherrschenden Geschlechter-Leistungsstereotypen entspricht (welches eine schlechtere Leistung von Jungen im Fach Deutsch beinhaltet, dazu kongruenter Leistungsabfall bei bedrohten Jungen). Zusätzlich zeigte sich eine positive Auswirkung einer Darstellung von Jungen als generelle Bildungsverlierer auf Mädchen im Sinne eines *stereotype lift*-Effekts ebenfalls nur domänenspezifisch in dem Fach (Deutsch), in welchem das Leistungsstereotyp für die bessere Leistung von Mädchen spricht. Da insbesondere in Mathematik, einem sog. „Jungen-Fach“, keine negativen Effekte für Jungen in Studie 2 festgestellt werden konnten, soll überprüft werden, ob Jungen sich als Reaktion auf die Darstellung als Schulversager reaktiv verstärkt einem Bereich zuwenden, in welchem die stereotypen Erwartungen für ihre Leistung sprechen.

7.2. Methode

Teilnehmende. Für diese Untersuchung konnte eine Stichprobe von insgesamt 100 Schüler/-innen eines Gymnasiums im Berliner Außenbezirk Spandau gewonnen werden, hierauf entfallen 49 Jungen und 51 Mädchen. Erneut mussten Probandinnen und Probanden aus den weiterführenden Analysen ausgeschlossen werden, da die Testhefte nur in geringstem Maße ausgefüllt (<25% der Testhefte wurden bearbeitet), Muster angegeben (diagonale Ankreuzmuster, ausschließliches Ankreuzen der mittleren Kategorie) oder mehr als 75% der Fragen zur Manipu-

lationsüberprüfung nicht bzw. falsch beantwortet wurden. Dies betrifft 8 Jungen und 5 Mädchen, was in einer bereinigten Gesamtstichprobe von $N = 87$ (41 Jungen, 46 Mädchen) resultiert. Diese Teilnehmer/-innen waren im Mittel $M = 14.45$ Jahre alt ($SD = 0.54$, $range = 14-16$ Jahre).

Für diese Untersuchung waren die gleichen Richtlinien und Vorgaben der Berliner Senatsverwaltung für Bildung und Forschung zu berücksichtigen, weshalb sich eine explizite Erfassung des Migrationshintergrundes nur auf die befragten Schüler/-innen direkt beziehen durfte. Insgesamt geben $n = 3$ Jungen und $n = 3$ Mädchen eine direkte Migrationserfahrung an, was wiederum einer geringen Quote von Migrantinnen und Migranten im Klassenverbund entspricht. Unter Verwendung der gesprochenen Sprache als Indikator für eine Migrationserfahrung in zweiter (oder höherer) Generation liegen Angaben von $n = 62$ Schüler/-innen vor. Hier geben $n = 6$ Jungen (20.7%) und $n = 5$ Mädchen (15.2%) an, mit den Eltern eine andere Sprache als Deutsch zu sprechen. Zwischen den Geschlechtern und ebenfalls zwischen den Untersuchungsbedingungen bestehen keine bedeutsamen Unterschiede hinsichtlich der Migrationserfahrung, ganz gleich welcher Indikator herangezogen wird, alle $\chi^2 < 1.53$, $p > .22$.

Ebenfalls kongruent zu den vorhergehenden Studien wurde der sozioökonomische Hintergrund mittels Fragen zu häuslichen Besitztümern („*home possessions*“, Frey et al. 2009) operationalisiert. Hier zeigt sich hinsichtlich der abgefragten häuslichen Besitztümer eine mittlere Verfügbarkeit von $M = 5.79$ ($SD = 0.96$, $range = 3 - 7$), was wiederum einen vergleichsweise hohen sozioökonomischen Status indiziert. Eine Analyse des sozioökonomischen Status mit einer univariaten ANOVA weist mit $F < 1.46$, $p > .23$ keine bedeutsamen Unterschiede zwischen den Geschlechtern ($M_{\text{Jungen}} = 5.76$, $SD = 1.07$; $M_{\text{Mädchen}} = 5.83$, $SD = 0.87$) oder den Untersuchungsbedingungen ($M_{\text{EG}} = 5.94$, $SD = 0.75$; $M_{\text{KG}} = 5.66$, $SD = 1.12$) i.S.v. signifikanten Haupt- und Interaktionseffekten auf. Ergänzend wurde auch in dieser Stichprobe das aktuelle Notenspektrum (im Sinne der letzten Zeugnisnote) erfasst. Im Mittel über die gesamte Stichprobe ergibt sich eine Leistung in Deutsch von $M = 2.82$ ($SD = 0.84$, $range = 1 - 5$), Mathematik $M = 2.99$ ($SD = 1.05$, $range = 1 - 6$), Gesamtnotendurchschnitt $M = 2.85$ ($SD = 0.59$, $range = 1.40 - 5.14$). Somit liegt sowohl in den zentralen Domänen für diese Untersuchung als auch bezüglich des Gesamtnotendurchschnitts eine im Mittel zufriedenstellende Leistung der untersuchten Schüler/-innen vor, welche das gesamte Notenspektrum abdeckt und damit für eine realitätsnahe Leistungsverteilung der untersuchten Stichprobe spricht. Zwischen Jungen und Mädchen unterscheiden sich die Leistungen, resp. die Noten nicht statistisch bedeutsam; Deutsch: $M_{\text{Jungen}} = 2.90$, $M_{\text{Mädchen}} = 2.76$; Mathematik: $M_{\text{Jungen}} = 2.84$, $M_{\text{Mädchen}} = 3.12$; Durchschnittsnote: $M_{\text{Jungen}} = 2.81$, $M_{\text{Mädchen}} = 2.89$; alle $t < |1.20|$, $p > .24$.

Die Probandinnen und Probanden wurden zufällig einer Experimental- (*stereotype threat*) oder Kontrollbedingung (*no threat*) zugewiesen. Hier entfallen $n = 46$ auf die Experimentalgrup-

pe (20 Jungen und 26 Mädchen) sowie $n = 41$ auf die Kontrollgruppe (21 Jungen und 20 Mädchen). Zwischen den Gruppen wurden hinsichtlich der Schulleistungen (Noten) keine bedeutsamen Unterschiede festgestellt. Eine Übersicht über die Ergebnisse liefert Tabelle 6:

Tabelle 6: Einzelvergleiche aktueller Deutsch-, Mathematik- und Durchschnittsnote in Abhängigkeit der Untersuchungsbedingung. Zu $N = 87$, resp. $df = 85$ fehlende Fälle entsprechen der Anzahl fehlender Angaben.

Variable	Mittelwerte	T-Wert	p
	Untersuchungsbed.	(df)	
Deutschnote letztes Zeugnis	EG $M = 2.73$	-0.99 (75)	.33
	KG $M = 2.92$		
Mathematiknote letztes Zeugnis	EG $M = 3.05$	0.54 (76)	.59
	KG $M = 2.92$		
Durchschnittsnote letztes Zeugnis	EG $M = 2.89$	0.57 (76)	.57
	KG $M = 2.81$		

Hinsichtlich der initial erhobenen Merkmale und Leistungsdaten verteilen sich alle erfassten und untersuchten Variablen sowohl gleichsam auf die beiden Geschlechtergruppen als auch auf die Experimental- und Kontrollbedingung, was für eine hohe Vergleichbarkeit dieser Gruppen spricht. Somit wurde eine sichere Basis für die geplante Manipulation geschaffen, d.h. es kann ausgeschlossen werden, dass evtl. aufzufindende Effekte aufgrund der Heterogenität der untersuchten Gruppen zustande kommen und nicht auf die eigentliche, experimentelle Manipulation zurückgeführt werden können.

Prozedur und Material. Die vorliegende Studie wurde in allen untersuchten Klassen zu gleichen Bedingungen, in zwei aufeinanderfolgenden Schulstunden am Mittag im regulären Klassenverband erhoben. Einführend wurde nach der Begrüßung der Teilnehmenden eine mittels Untersuchungsleitfaden standardisierte Instruktion vorgetragen, welche die Schüler/-innen mit der Arbeit mit dem Testheft vertraut machten. Zusätzlich wurde zur Erhöhung der *test-compliance* eine Entlohnung von 10 Euro je ernsthaft bearbeiteten Fragebogen für die Klassenkasse in Aussicht gestellt.

Hiernach begann die eigentliche Arbeit mit dem Testheft unter Zusicherung vollständiger Anonymität. Um neben den Leistungsunterschieden im Sinne von Noten die initiale Leistung der Schüler/-innen kontrollieren zu können und die Messung der motivationalen Zielorientierung rechtfertigen zu können, wurde ein Leistungstest, vergleichbar mit dem in Studie 2, zur Bearbeitung vorgelegt. Dieser bestand aus 15 Aufgaben zum Leseverständnis und 12 Aufgaben aus dem mathematischen Bereich aus unveröffentlichten Aufgaben aus VERA-8 (Institut zur Qualitätsentwicklung im Bildungswesen, 2008). Die Reihenfolge, ob zuerst Aufgaben aus dem Fach Deutsch oder aus der Mathematik zu bearbeiten sind, wurde ausbalanciert. Da diese Untersuchung an der höchsten deutschen Schulform, dem Gymnasium durchgeführt wurde, wurde die

Aufgabenschwierigkeit der zu bearbeitenden Aufgaben geringfügig erhöht. Nach dem Leistungstest wurde eine kurze Pause gewährt, nach welcher das experimentelle Treatment appliziert wurde. Im Gegensatz zu Studie 2 war hier kein reiner Lesetext mehr zu bearbeiten (auch, um den Kompetenznachteilen von Jungen im Lesen Rechnung zu tragen, OECD, 2014), vielmehr wurden die entsprechenden Informationen in Tabellen und Grafiken umgesetzt, zu denen nur noch ein kurzer, zusammenfassender und erläuternder Text präsentiert wurde. Zusätzlich diente diese Erneuerung des Treatments einer Überprüfung, ob die negativen Effekte, wie in Studie 2 gezeigt, von den direkten, expliziten Formulierungen des Textes abhängig sind oder auch auf subtilere Art und Weise zum Tragen kommen. In der Experimentalbedingung wurden eine Tabelle und zwei Abbildungen vorgelegt, aus denen im Zuge der Dateninterpretation ersichtlich wurde, dass Jungen vergleichsweise häufig geringere Schulformen besuchen als Mädchen und dementsprechend auch geringer wertige Abschlusszertifikate (entnommen aus Mikrozensus, Statistisches Bundesamt, 2014a) und weniger Kompetenzen erwerben. In der Kontrollbedingung wurden vergleichbare Materialien (Tabelle und Abbildungen) verwendet, allerdings ohne jedwede Bezugnahme auf Geschlechterunterschiede, sondern mit einem Fokus auf sich unabhängig vom Geschlecht ändernde Schulbesuchszahlen auf verschiedenen Schulformen. In beiden Bedingungen wurden darauf folgend kongruent zu Studie 2 in der Experimentalgruppe sechs Aufgaben (Aufgabe 4 bestand hier aus 3 *single-choice* Unteraufgaben), in der Kontrollgruppe vier Aufgaben geringer Schwierigkeit zur Manipulationsüberprüfung bearbeitet, um sicherzustellen, dass alle notwendigen Informationen aus dem Material entnommen wurden.

Darauf folgend wurde angekündigt, dass noch ausreichend Zeit vorliegt, um weitere Aufgaben aus den Bereichen Deutsch und Mathematik – ähnlich den zu Beginn bearbeiteten Aufgaben – zu bearbeiten. Unter dem Vorwand, die individuellen Bedürfnisse bei der Aufgabenbearbeitung hier aber mehr berücksichtigen zu wollen, wurde in ausbalancierter Reihenfolge für beide untersuchten Domänen, Deutsch und Mathematik, eine 1-Item-Messung der motivationalen Zielorientierung (Leggett, 1985) vorgelegt. Die Probanden/-innen wurden gebeten anzugeben, welche Art von Aufgabe sie im Folgenden in der entsprechenden Domäne bearbeiten möchten: „Aufgaben, die nicht zu schwierig sind, damit ich wenig Fehler mache“ (*performance avoidance goal*), „Aufgaben, die sehr einfach sind, sodass ich diese gut lösen kann“ (*performance avoidance goal*), „Aufgaben, die ziemlich einfach sind, um zeigen zu können, was ich bereits alles kann“ (*performance approach goal*) sowie „Aufgaben, von denen ich viel lernen kann, auch wenn ich dabei nicht besonders klug aussehe“ (*learning goal*).

Abschließend wurden die Teilnehmenden informiert, dass keine weiteren Aufgaben zu bearbeiten sind und gebeten, die soziodemografischen Fragen zu Alter, Geschlecht usw. zu beantworten. Nach der Bearbeitung dieser letzten Fragen erfolgte eine vollständige Aufklärung über Inhalt und Ziel der Untersuchung sowie die Überreichung eines symbolischen Schecks für

die in Aussicht gestellte Entlohnung. Insgesamt dauerte die gesamte Untersuchung ca. 60 (\pm 5) Minuten.

7.3. Ergebnisse

Berechnung der Leistungsdaten und initiale Leistung im Pretest. Zur Berechnung des initialen Leistungsniveaus der Schüler/-innen bezüglich der Deutsch- und Mathematikaufgaben wurden wie in Studie 2 die korrekt beantworteten Fragen gezählt und in ein Verhältnis zur Gesamtzahl der Fragen gesetzt, nicht beantwortete Fragen wurden als falsch gewertet. Dieser Indikator für die Leistung in beiden erfragten Domänen wurde einer ANOVA unterzogen, um Unterschiede zwischen den Geschlechtern oder innerhalb der Untersuchungsgruppen feststellen zu können. Es konnten keine bedeutsamen Leistungsunterschiede in Deutsch oder Mathematik zwischen der Experimental- und Kontrollbedingung festgestellt werden; Leseaufgaben: $M_{EG} = 63.40\%$, $SD = 20.23\%$, $M_{KG} = 63.17\%$, $SD = 18.43\%$; $F < 1$, n.s.; Mathematikaufgaben: $M_{EG} = 56.60\%$, $SD = 19.59\%$, $M_{KG} = 54.50\%$, $SD = 19.34\%$; all $F < 1$, n.s. Weiterhin unterschied sich die initiale Leistung auch nicht zwischen den Geschlechtern; Leseaufgaben: $M_{Jungen} = 62.50\%$, $SD = 21.06\%$; $M_{Mädchen} = 63.97\%$, $SD = 17.89\%$; $F < 1$, n.s.; Mathematikaufgaben: $M_{Jungen} = 58.00\%$, $SD = 19.11\%$; $M_{Mädchen} = 53.61\%$, $SD = 19.61\%$; $F < 1$, n.s. (beide Interaktionen Experimentalbedingung \times Geschlecht, $F < 1$).

Hinsichtlich der Fragen zur Manipulationsüberprüfung wurden nur Schüler/-innen in die weiteren Analysen einbezogen, welche mindestens 50% der entsprechenden Fragen korrekt beantwortet haben (was aufgrund des sich wiederholenden Charakters der inhaltlichen Aussagen innerhalb der Kontrollfragen als ausreichend erachtet wurde). Alle 87 Teilnehmer/-innen haben dieses Kriterium erfüllt, die wie oben beschrieben ausgeschlossenen Schüler/-innen (8 Jungen und 5 Mädchen) haben dieses Kriterium nicht erreicht und wiesen weitere Auffälligkeiten auf (z.B. angekreuzte Muster, erhöhtes Aufkommen fehlender Werte), weshalb diese nicht in die zu analysierende Stichprobe aufgenommen werden konnten. Insgesamt liegt mit $M = 65.23\%$ ($SD = 14.16\%$, $range = 50 - 100\%$) ein zufriedenstellendes Ergebnis hinsichtlich der Fragen zur Manipulationsüberprüfung vor, welches indiziert, dass ein substanzieller Anteil des Materials zur Manipulation angemessen verarbeitet und verstanden wurde. Eine ANOVA weist keine bedeutsamen Unterschiede zwischen den Untersuchungsbedingungen oder den Geschlechtern (Haupt- und Interaktionseffekte: $F < 1.48$, $p > .23$) auf. In der Kontrollgruppe wurde ein geringfügig höherer Anteil der Kontrollfragen richtig beantwortet, $M_{EG} = 64.13\%$ ($SD = 14.03\%$); $M_{KG} = 66.46\%$ ($SD = 14.37\%$), welcher mit $t(85) < |1|$, $p = .45$ allerdings nicht signifikant ausfällt. Dies deutet auf eine vergleichbare Aufgabenschwierigkeit der Fragen zur Manipulationsüberprüfung zwischen beiden Treatment-Bedingungen hin. Zwischen Jungen ($M = 65.04\%$, $SD = 15.72\%$) und

Mädchen ($M = 65.40\%$, $SD = 12.78\%$) lassen sich mit $t(85) < |1|$, $p = .91$ ebenfalls keine divergierenden Leistungen in Bezug auf die Kontrollfragen feststellen.

Analyse der motivationalen Daten. Die zwei Items, welche ein Vermeidungs-Leistungsziel beschreiben, wurden Dweck & Leggett (1988) folgend in eine gemeinsame Kategorie zusammengefasst. Im Folgenden wurden getrennt nach den Fächern Deutsch und Mathematik die Häufigkeiten der durch die Schüler/-innen angegebenen Zielorientierungen getrennt nach Geschlecht und der Untersuchungsbedingung ausgezählt. Einen Überblick über die Zielorientierung im Fach Deutsch gibt Tabelle 7:

Tabelle 7: Chi-Quadrat-Analysen und Häufigkeitsverteilung der motivationalen Zielorientierung im Fach Deutsch in Abhängigkeit von Geschlecht und Untersuchungsbedingung.

	Jungen			Mädchen		
	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel
Experimental- gruppe	5	5	10	5	4	13
Kontrollgruppe	4	7	9	3	8	8
	$\chi^2(2, N = 40) = .50, p = .78$			$\chi^2(2, N = 41) = 2.82, p = .24$		

Die Ergebnisse im Fach Deutsch zeigen, dass in etwa die Hälfte der befragten Schüler/-innen (19 Jungen und 21 Mädchen) eine Lernzielorientierung aufweisen. Die verbleibenden 21 Jungen und 20 Mädchen weisen eine Leistungszielorientierung auf, hiervon wiederum ca. die Hälfte eine Vermeidungs-Leistungszielorientierung (9 Jungen und 8 Mädchen) bzw. eine Annäherungs-Leistungszielorientierung (12 Jungen, 12 Mädchen). Beim Vergleich dieser Ergebnisse zwischen den Gruppen wird deutlich, dass hinsichtlich der antizipierten, zusätzlichen Leseaufgaben keine bedeutsamen Unterschiede bei Jungen ($\chi^2(2, N = 40) = .50, p = .78$) und Mädchen ($\chi^2(2, N = 41) = 2.82, p = .24$) in Abhängigkeit von der Untersuchungsbedingung auftreten. Circa die Hälfte der Mädchen und ein ebenso großer Anteil der Jungen verfolgt eine Lernzielorientierung, unabhängig vom applizierten Treatment.

In der Domäne Mathematik zeigt sich hingegen ein abweichendes Ergebnismuster, welches in Tabelle 8 zusammenfassend dargestellt ist:

Tabelle 8: Chi-Quadrat-Analysen und Häufigkeitsverteilung der motivationalen Zielorientierung im Fach Mathematik in Abhängigkeit von Geschlecht und Untersuchungsbedingung.

	Jungen			Mädchen		
	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel
Experimental- gruppe	5	2	12	11	3	8
Kontrollgruppe	10	8	3	10	5	4
	$\chi^2(2, N = 40) = 10.59, p < .01$			$\chi^2(2, N = 41) = 1.67, p = .43$		

Der Großteil der Schüler/-innen verfolgt in Mathematik ein Vermeidungs-Leistungsziel (25 Jungen und 29 Mädchen), wobei ein substanzieller Anteil dieser Gruppe eine Vermeidungs-Leistungszielorientierung einnimmt (15 Jungen und 21 Mädchen) und nur 10 Jungen und 8 Mädchen ein Annäherungs-Leistungsziel aufweisen. Nur 15 Jungen und 12 Mädchen gaben eine Lernzielorientierung bei den mathematischen Aufgaben an. Beim Vergleich dieser Ergebnisse unter Berücksichtigung der Experimentalbedingung fällt allerdings auf, dass Jungen nach der Bearbeitung des Treatments über den geringeren Schulerfolg von Jungen wesentlich häufiger eine Lernzielorientierung berichteten und wesentlich weniger häufig eine Annäherungs- oder Vermeidungs-Leistungszielorientierung einnahmen als Jungen in der Kontrollbedingung, $\chi^2(2, N = 40) = 10.59, p < .01$. Werden Vermeidungs- und Annäherungs-Leistungsziele in eine gemeinsame Oberkategorie „Leistungszielorientierung“ zusammengefasst, ergibt sich eine noch deutlichere statistische Bedeutsamkeit der beschriebenen Unterschiede, $\chi^2(1, N = 40) = 10.17, p = .001$. Der Effekt bleibt ebenfalls stabil, wenn nur Vermeidungs-Leistungsziele als Vergleichsdimension herangezogen werden, $\chi^2(1, N = 30) = 6.65, p < .01$. Konträr dazu konnte für Mädchen kein Effekt der Experimentalbedingung festgestellt werden, $\chi^2(2, N = 40) = 1.67, p = .43$.

Um diese Ergebnisse mit einer einzelnen Analyse absichern zu können und um einen möglichen Einfluss der erhobenen Leistungsdaten kontrollieren zu können wurde eine binärlogistische Regressionsanalyse durchgeführt (AV: motivationale Zielorientierung, 1 = Lernzielorientierung). Hierzu wurde die motivationale Zielorientierung in eine Dummy-Variable überführt (0 = *performance goal*, 1 = *learning goal*), als Prädiktoren dienten Geschlecht (1 = männlich), die Experimentalbedingung (1 = EG), Leistungsdomäne (1 = Mathematik), die Interaktionsterme aus Geschlecht \times Experimentalbedingung sowie der Dreifachinteraktionsterm Geschlecht \times Experimentalbedingung \times Leistungsdomäne und die Leistungsscores in Deutsch und Mathematik. Die Ergebnisse unterstützen die Ergebnismuster, welche in den Chi-Quadrat Analysen berichtet wurden: die Leistungsdomäne entspricht mit $B = -1.16, SE(B) = .404, Wald = 8.19, p < .01$ einem signifikanten Prädiktor, welcher aber durch die ebenfalls signifikante Dreifachinteraktion bedingt wird, $B = 1.69, SE(B) = .766, Wald = 4.85, p < .05, R^2 = .14$. Die Richtung der Dreifachinteraktion gibt an, dass Jungen in der Experimentalbedingung nur im Fach Mathematik wahrscheinlicher eine Lernzielorientierung einnehmen, nicht jedoch im Fach Deutsch und auch

ohne signifikanten Effekt für Mädchen. Zusätzlich ist anzumerken, dass die aktuelle Leistung in Deutsch oder Mathematik keinen prädiktiven Wert für motivationale Zielorientierungen aufwies.

Da dieses Ergebnismuster nur zum Teil den Vorhersagen der formulierten Hypothesen hinsichtlich der motivationalen Reaktion von Jungen auf die Exposition der Stereotypenbedrohung von schulversagenden Jungen entspricht, wurde zur Absicherung des Befundes, dass Jungen auf eine Stereotypenbedrohung motivational nur mit einer Steigerung ihrer Lernzielorientierung in Mathematik reagieren, eine Replikation durchgeführt. Die Diskussion beider Befunde erfolgt gemeinsam nach Darstellung der Ergebnisse der replizierenden Studie.

7.4. Studie 4: Replikation

Teilnehmende. Für eine Replikation der Untersuchung zu den Auswirkungen von *stereotype threat* auf die motivationalen Zielorientierungen von Jungen wurden nochmals 88 Probanden/-innen einer Berliner Sekundarschule im Bezirk Mitte gewonnen, hiervon 42 Jungen und 46 Mädchen. Entsprechend dem Vorgehen in Studie 3 mussten erneut einige Versuchsteilnehmer/-innen (je 4 Jungen und Mädchen) von den weiterführenden Analysen ausgeschlossen werden, wobei exakt gleiche Bedingungen wie in Studie 3 zugrunde gelegt wurden. Die verbleibenden $N = 80$ Schülerinnen und Schüler waren durchschnittlich $M = 14.60$ Jahre alt ($SD = 0.71$ Jahre, *range* 14 – 17 Jahre). Drei Jungen und 5 Mädchen wiesen eine direkte Migrationserfahrung auf, wird die gesprochene Sprache mit den Eltern als zusätzlicher Indikator hinzugezogen sprachen $n = 6$ Jungen und $n = 5$ Mädchen mit ihren Eltern eine andere Sprache, als Deutsch⁹. Zwischen den Geschlechtern oder den Untersuchungsbedingungen konnten hinsichtlich des Migrationshintergrundes oder der gesprochenen Sprache keine bedeutsamen Unterschiede festgestellt werden, $\chi^2 < 3.66$, $p > .16$. Der sozioökonomische Status lag mit $M = 5.00$ ($SD = 1.22$, *range* 2 – 7) etwas unter dem Mittelwert des sozioökonomischen Status („*home possessions*“) aus Studie 3, divergiert aber erneut nicht nach Geschlecht, $M_{Jungen} = 4.95$ ($SD = 1.29$), $M_{Mädchen} = 5.05$ ($SD = 1.13$), $F < 1$, n.s. Zwischen den Untersuchungsbedingungen zeigt sich, dass innerhalb der Experimentalgruppe die Verfügbarkeit der abgefragten Güter geringfügig höher ausfällt, $M_{EG} = 5.23$ ($SD = 1.23$), $M_{KG} = 4.67$ ($SD = 1.24$), $F(1,75) = 4.01$, $p = .05$, $\eta^2 = .05$. Da beide Werte aber dem oberen Durchschnitt vergleichbarer Erhebungen (OECD, 2011) entsprechen, wird diese geringfügige Differenz als unproblematisch erachtet. Die Interaktion fällt mit $F < 1$ nicht statistisch bedeutsam aus.

⁹Zusätzlich sprechen $n = 6$ Jungen und $n = 9$ Mädchen mit den Eltern sowohl Deutsch als auch eine andere Sprache.

Prozedur und Material. Die Teilnehmenden wurden kongruent zu Studie 3 randomisiert auf die Experimental- oder Kontrollbedingung zugewiesen, wobei 41 Personen (20 Jungen und 21 Mädchen) auf die Experimentalbedingung und 39 Teilnehmer/-innen (18 Jungen und 21 Mädchen) auf die Kontrollgruppe entfielen. Die Prozedur entsprach vollständig der Umsetzung in Studie 3 mit der Ausnahme, dass aufgrund zeitlicher Restriktionen und fehlender Hinweise auf einen Einfluss der aktuellen Leistung (siehe *Studie 3*) keine Leistungsmessung stattfand und keine aktuellen Leistungen im Sinne von Schulnoten erhoben wurden. Nach der Bearbeitung einer Übungsaufgabe und der Applikation des experimentellen Treatments incl. der Beantwortung der Fragen zur Manipulationsüberprüfung wurde den Schüler/-innen stattdessen jeweils eine Beispielaufgabe aus den Bereichen Deutsch und Mathematik (aus dem Pretest zu Studie 3) vorgelegt, gefolgt von der Instruktion, dass nunmehr Leistungsaufgaben zu bearbeiten sind und der darauf aufbauenden Frage „*Welche Aufgaben möchtest du im Folgenden bearbeiten?*“. Daraufhin wurde die motivationale Zielorientierung kongruent zur vorherigen Studie mit dem Messinstrument von Leggett (1985), d.h. mit der Vorlage von vier Antwortoptionen, die jeweils verschiedene Zielorientierungen erfassen, erfragt. Die Gesamtdauer dieser Untersuchung betrug ca. 45 Minuten.

Ergebnisse. Die Analyse der Fragen zur Manipulationsüberprüfung zeigte, dass insgesamt $M = 60.00\%$ ($SD = 12.95\%$, *range* 50 – 100%) der Fragen zur Manipulationsüberprüfung richtig beantwortet wurden, womit die untersuchte Stichprobe (nach Ausschluss der o.g. 8 Probanden/-innen) die initiale Anforderung von mindestens 50% korrekter Antworten, um in die zu analysierende Stichprobe aufgenommen zu werden, erfüllte. In der Experimentalgruppe wurden im Mittel $M_{EG} = 60.98\%$ ($SD = 13.75\%$, *range* = 50 – 100%), in der Kontrollbedingung $M_{KG} = 58.97\%$ ($SD = 12.15\%$, *range* = 50 – 100%) richtig beantwortet. Aufgrund des sich wiederholenden, inhaltlichen Charakters dieser Fragen ist somit ein ausreichendes Leistungsniveau bezüglich der Kontrollfragen festzustellen, welches im Vergleich zu den Ergebnissen der *Studie 3* zwar etwas geringer ausfällt, aber als noch vergleichbar anzusehen ist. Die Leistung in den Fragen zur Manipulationsüberprüfung wurde mit einer ANOVA auf Unterschiede innerhalb der Gruppen (Geschlecht und Experimentalbedingung) überprüft, welche keine signifikanten Haupt- und Interaktionseffekte ergab ($F < 1$, $p > .41$). Somit divergierten die Anteile korrekt beantworteter Fragen zur Manipulationsüberprüfung nicht zwischen den Untersuchungsgruppen, $t(78) < 1$, $p = .52$, was die bisherigen Befunde über eine vergleichbare Schwierigkeit der Untersuchungsmaterialien einschließlich Kontrollfragen unterstützt. Ebenso konnte kein bedeutsamer Unterschied zwischen Jungen und Mädchen festgestellt werden ($M_{\text{Jungen}} = 60.53\%$, $SD = 13.79\%$; $M_{\text{Mädchen}} = 59.52\%$, $SD = 12.29\%$; $t(78) < 1$, $p = .73$).

Bezüglich der erhobenen Angaben zur motivationalen Zielorientierung im Fach Deutsch zeigte sich erneut, dass Jungen der Experimentalbedingung in dieser Domäne nach der Bearbei-

tung des bedrohenden Materials („Jungen als Schulversager“) nicht von den motivationalen Zielorientierungen der Probanden in der Kontrollgruppe abweichen:

	Jungen			Mädchen		
	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel
Experimental- gruppe	8	7	5	6	8	6
Kontrollgruppe	9	5	4	7	5	8
	$\chi^2 (2, N = 38) = 0.40, p = .82$			$\chi^2 (2, N = 40) = 1.06, p = .59$		

In beiden Untersuchungsbedingungen stellte für Jungen eine Leistungszielorientierung die am meisten präferierte Option dar. In der Experimentalgruppe entschieden sich 15 jungen für diese motivationale Zielorientierung – in der Kontrollgruppe lag mit 14 Personen, welche diese Zielorientierung präferieren, eine vergleichbare Häufigkeit vor. Auch bei Differenzierung der Leistungszielorientierung in eine Annäherungs- und Vermeidungskomponente ergaben sich keinerlei bedeutsame Unterschiede in den Häufigkeiten der Angaben zwischen den Gruppen, ebenso wenig bei der Lernzielorientierung, $\chi^2 (2, N = 38) = 0.40, p = .82$. Auch Mädchen weichten nicht bedeutsam von diesem Befundmuster ab, die Leistungszielorientierung war hier mit jeweils 13 Wahlen ebenfalls die am meisten präferierte Option, eine Differenzierung nach Annäherungs- und Vermeidungsmotiv ergab ebenfalls keinen Hinweis auf bedeutsame Unterschiede in der motivationalen Zielorientierung. Auch bezugnehmend auf die Lernzielorientierung wiesen die Häufigkeiten nicht auf bedeutsame Unterschiede hin, was die statistische Analyse insgesamt bestätigt, $\chi^2 (2, N = 40) = 1.06, p = .59$. Im Fach Deutsch konnte somit weder für Jungen, noch für Mädchen ein Einfluss des experimentellen Treatments (Jungen als Bildungsverlierer vs. Wandel des deutschen Schulsystems ohne jedwede Bezugnahme auf Geschlecht) auf die motivationale Zielorientierung festgestellt werden.

Im Fach Mathematik zeigte sich wiederum ein anderes Ergebnismuster, wie die Häufigkeitsverteilungen und -analysen in der folgenden Tabelle zusammenfassend illustriert:

Tabelle 10: Chi-Quadrat-Analysen und Häufigkeitsverteilung der motivationalen Zielorientierung im Fach Mathematik in Abhängigkeit von Geschlecht und Untersuchungsbedingung.

	Jungen			Mädchen		
	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel	Vermeidungs- Leistungsziel	Annäherungs- Leistungsziel	Lernziel
Experimental- gruppe	4	4	12	10	4	7
Kontrollgruppe	11	5	2	13	3	5
	$\chi^2(2, N = 38) = 10.44, p < .01$			$\chi^2(2, N = 42) = 0.87, p = .70$		

Während Jungen in der Experimentalbedingung mit $n = 12$ am häufigsten eine Lernzielorientierung angaben, nahm ein Großteil ($n = 16$) von Jungen der Kontrollgruppe eine Leistungszielorientierung ein – hiervon entfiel wiederum ein substantieller Anteil ($n = 11$) auf eine Vermeidungs-Leistungszielorientierung, $\chi^2(2, N = 38) = 10.44, p < .01$. Diese Unterschiede zwischen Jungen der Experimental- und Kontrollgruppe bleiben stabil, wenn Vermeidungs- und Annäherungs-Leistungsziele in eine gemeinsame Kategorie aggregiert werden, $\chi^2(1, N = 38) = 9.73, p < .01$ oder wenn die Lernzielorientierung ausschließlich mit der Vermeidungs-Leistungszielorientierung kontrastiert wird, $\chi^2(1, N = 29) = 10.20, p < .001$. Bei Mädchen zeigten sich keine bedeutsamen Häufigkeitsunterschiede hinsichtlich der motivationalen Zielorientierung in Mathematik: Unabhängig von der experimentellen Bedingung ($\chi^2(2, N = 42) = 0.87, p = .70$) dominierte bei Mädchen die Leistungszielorientierung (EG: $n = 14$, KG: $n = 16$) mit einem Schwerpunkt auf einer Vermeidungs-Leistungszielorientierung.

Wie in Studie 3 wurden unter Zuhilfenahme einer binär-logistischen Regressionsanalyse die soeben präsentierten, getrennten Häufigkeitsanalysen in Deutsch und Mathematik in einer Einzelanalyse zusammengefasst. Als abhängige Variable diente dabei die motivationale Zielorientierung (1 = Lernzielorientierung). Als Prädiktoren wurden Geschlecht (1 = männlich), Untersuchungsbedingung (1 = Experimentalgruppe), die untersuchte Domäne (1 = Mathematik) sowie der Interaktionsterm Geschlecht \times Untersuchungsbedingung und die Tripel-Interaktion Geschlecht \times Untersuchungsbedingung \times Fach/Domäne eingeführt. In diesem Regressionsmodell zeigte sich, dass lediglich die Dreifachinteraktion als ein bedeutsamer Prädiktor hervortrat, $B = 1.91, SE(B) = .808, Wald = 5.60, p < .05, R^2 = .11$. Keine anderen Prädiktoren erreichten eine statistische Bedeutsamkeit ($B < |.853|, p > .13$). Dieses Ergebnis verdeutlicht, dass ausschließlich Jungen der Experimentalgruppe im Fach Mathematik im Vergleich zu Jungen der Kontrollgruppe wahrscheinlicher eine Lernzielorientierung einnehmen, wohingegen keine Unterschiede zwischen den Untersuchungsbedingungen im Fach Deutsch aufzufinden sind.

7.5. Gesamtdiskussion

Ziel der vorliegenden Untersuchungen war es festzustellen, welche Auswirkungen eine generelle Stereotypenbedrohung von Jungen jenseits von Effekten auf die (schulische) Leistungsfähigkeit aufweist, um ferner einen Erklärungsansatz zu erarbeiten, inwiefern das Paradigma *stereotype threat* den vergleichsweise überdauernden, geringeren Schulerfolg von Jungen erklären kann. Hierzu wurde versucht, einen psychologischen Prozess hinter dem Phänomen der Stereotypenbedrohung, der motivationalen Zielorientierung, zu beschreiben (Ryan & Ryan, 2005).

Die Ergebnisse beider Studien verdeutlichen, dass Jungen nach der Exposition von Informationen über ihr angebliches Scheitern in der Schule (kongruent zur öffentlichen, medialen Berichterstattung) eine stärkere Lernzielorientierung im Fach Mathematik aufwiesen, als es bei Jungen nach der Bearbeitung von geschlechtsneutralem, schulbezogenem Material der Fall war. Im Fach Deutsch konnten keine bedeutsamen Unterschiede hinsichtlich der motivationalen Zielorientierung festgestellt werden, ebenso wurden Mädchen durch das experimentelle Treatment nicht beeinflusst. Somit liegt erneut ein domänenspezifischer Effekt vor, allerdings in anderer Art und Weise als in der vorhergehenden Untersuchung zum Einfluss einer Stereotypenbedrohung auf die schulische Performanz. Die Ergebnisse sind inkompatibel mit der Annahme, dass Jungen nach einer generellen Stereotypenbedrohung in dem Fach, in welchem die existierenden Geschlechter-Leistungstereotype gegen ihren Erfolg sprechen, eher eine lernhinderliche (Vermeidungs-) Leistungszielorientierung einnehmen, was zur Folge hat, dass Hypothese 6 abzulehnen ist. Hypothese 7 konnte bestätigt werden: Jungen der Experimentalbedingung wiesen häufiger eine Lernzielorientierung in einem für sie positiv leistungsstereotypisierten Fach auf als Jungen der Kontrollgruppe, was mit verstärkt lernförderlichen Verhaltensweisen, Expertise und Kompetenzerwerb einhergeht (Dweck & Leggett, 1988; Elliot, 1999; Huang, 2012). Da für Mädchen keinerlei bedeutsame Unterschiede zwischen Experimental- und Kontrollbedingung in beiden untersuchten Domänen aufzufinden waren, ist die Hypothese 8 abzulehnen. Im Kontrast zu den Ergebnissen aus Studie 2 profitierten Mädchen in dieser Untersuchung nicht von der Darstellung von Jungen als unterlegene Gruppe in der Schule, insbesondere nicht im Fach Deutsch, in welchem in der vorherigen Studie ein Leistungsvorteil von Mädchen in der Experimentalbedingung aufgezeigt werden konnte. Somit scheint es so zu sein, dass die Stereotypenbedrohung bei Mädchen nicht zu einer Änderung der motivationalen Zielorientierung führt. Im Fach Mathematik zeigten sich bei Mädchen ebenfalls keine bedeutsamen Unterschiede zwischen den experimentellen Bedingungen, wovon aufgrund der männlichen Konnotation dieses Faches und insbesondere unter Bezugnahme der Ergebnisse aus Studie 2 nicht ausgegangen werden konnte.

Die gefundenen Ergebnisse sprechen gegen eine Beeinträchtigung der motivationalen Zielorientierung von Jungen durch die applizierte Stereotypenbedrohung, da die gefunden motivationalen Effekte ausschließlich positiver, lernförderlicher Natur waren und ergänzen somit die bisher vorliegende Befundlage, welche sich ausschließlich auf die Nachteile (zumeist von Mädchen) in negativ stereotypisierten Leistungsdomänen konzentrierten (u.a. Smith, 2004, 2006; Smith et al., 2007; Brown & Josephs, 1999). Es scheint unter Bezugnahme auf das Ergebnismuster in der vorliegenden Studie vielmehr so zu sein, dass Jungen die per Selbstbericht erhobene motivationale Zielorientierung im Sinne einer *stereotype reactance* als kompensatorische Strategie nutzen, um sich gegen die überzeichnete Darstellung von Jungen als *generelle* Verlierer in der Schule zur Wehr zu setzen, da diese Darstellung offensichtlich nicht ihr alltägliches Erleben in der Schule angemessen widerspiegelt. Diese Interpretation kann auch zur Erklärung beitragen, warum Mädchen in keiner der untersuchten Domänen substanzielle Unterschiede hinsichtlich der motivationalen Zielorientierung berichtet haben. Wenn tatsächlich der Ausdruck einer bestimmten motivationalen Zielorientierung als kompensatorische Strategie angewendet wird, besteht für Mädchen hier kein Bedarf, im Sinne einer *stereotype reactance* zu agieren, da vom applizierten Material keine Bedrohung für die eigene, soziale Gruppe ausgeht und somit kein entsprechender Handlungsbedarf als notwendige Reaktion auf die Darstellung abgeleitet wird – vielmehr scheint es für Mädchen adäquat, den *status quo* beibehalten zu wollen (Molden & Finkel, 2010).

Der Vorteil der vorliegenden Studie besteht ferner darin, dass einerseits nicht ausschließlich auf die leistungshemmende Leistungszielorientierung fokussiert wurde und andererseits neben einer für Jungen negativ stereotypisierten Leistungsdomäne konträr eine weitere Domäne zum Untersuchungsgegenstand gemacht wurde, in welcher ein positives Leistungsstereotyp für die untersuchte soziale Gruppe vorliegt. Ebenfalls wurde durch den konsequenten Einbezug von Mädchen, d.h. Mitgliedern der Außengruppe, sichergestellt, dass die Ergebnisse für Jungen nicht isoliert betrachtet werden müssen, sondern im Hinblick auf den gemeinsamen Schulalltag beider Geschlechter die soziale Geschlechterstruktur im Lernkontext Schule möglichst angemessen berücksichtigt wird. Unter Bezugnahme auf die bisherigen Arbeiten zum Einfluss einer Stereotypenbedrohung auf die motivationale Zielorientierung weichen die Ergebnisse unter Berücksichtigung dieser Besonderheiten der vorliegenden Studie ab. Während u.a. Brown & Josephs (1999), Smith (2006), und Smith et al. (2007) nachweisen konnten, dass eine bedrohte soziale Gruppe von Mädchen auf die Exposition eines bedrohlichen Leistungsstereotyps mit einer verstärkten Einnahme einer Vermeidungs-Leistungszielorientierung in dieser bedrohten Domäne reagieren, zeigte sich in der vorliegenden Untersuchung kein derartiger Effekt: Im Fach Deutsch, welches typischerweise nicht mit der Leistungsfähigkeit von Jungen assoziiert ist (Plante et al., 2009; Nosek et al., 2002) konnte keine verstärkte Präferenz von bedrohten Jungen

für Vermeidungs-Leistungsziele nachgewiesen werden. Dies spricht entweder für die Tatsache, dass die für Mädchen festgestellten Befunde entweder nicht vollständig auf den Kontext von „Jungen als Bildungsverlierer“ übertragen werden können, d.h. dass unter Umständen bei Jungen andere Mechanismen jenseits der motivationalen Zielorientierung angesprochen werden oder aber andererseits die festgestellten Ergebnisse den Besonderheiten des Versuchsplans geschuldet sind. Diese Interpretation legt nahe, dass Jungen in Reaktion auf die generelle Bedrohung eine Kompensationsleistung durch lernmotivationale Zuwendung zu einem Fach erbringen, in welchem keine Bedrohung oder gar ein Leistungsvorteil wahr- oder angenommen wird.

Entsprechend den Vorhersagen des Modells von Smith (2004) wird auf Basis der vorliegenden Befunde demnach eine Erweiterung notwendig. Das Modell geht nur von einem Einfluss einer kompetenzrelevanten Stereotypenbedrohung auf eine Vermeidungs-Leistungszielorientierung (*Pfad 1* im Modell) aus, auch die Erweiterung des initialen Modells von Smith (2004) durch Thoman et al. (2013) betrachtet die Auswirkung einer kontextspezifischen Salienz von *stereotype threat* ausschließlich auf die Leistungszielorientierung, wenngleich hier zumindest beide Komponenten (Annäherung und Vermeidung) einbezogen werden. Die vorliegenden Ergebnisse legen vielmehr nahe, Lern- und Leistungsmotivation als Gesamtkonstrukt zu begreifen (d.h. auch Lernzielorientierungen einzubeziehen) und im Kontext von *stereotype threat* und *stereotype reactance* zu betrachten. Dies impliziert auch, nicht nur die für die entsprechende soziale Gruppe negativ stereotypisierte Leistungsdomäne zu erfassen, sondern darüber hinaus auch Änderungen in der motivationalen Zielorientierung stärker zu berücksichtigen, welche jenseits der negativ stereotypisierten in anderen Leistungsdomänen liegen. Anders formuliert scheint es wichtig, eine Leistungssituation weniger abgegrenzt zu begreifen, sondern verstärkt darauf zu fokussieren, dass es auch Inhaltsbereiche jenseits der negativ stereotypisierten Leistungsdomäne gibt. Während bei alleiniger Zentrierung auf einen Fähigkeitsbereich evtl. eine Vermeidungs-Leistungszielorientierung vorherrschend ist, kann sich dieses Befundmuster unter Hinzunahme weiterer Leistungsbereiche durchaus im Sinne von *stereotype reactance* wandeln.

Zusätzlich wird die o.g. Interpretation durch Befunde von Seibt & Förster (2004) gestützt, welche zusätzlich auf die Möglichkeit hinweisen, dass die per Selbstbericht gemachten Angaben zur motivationalen Zielorientierung nicht nur das Bedürfnis widerspiegeln, das generell negative Leistungsstereotyp gegenüber Jungen zu entkräften, sondern auch dass via *stereotype reactance* tatsächlich das Bestreben zur Bearbeitung schwieriger, lernförderlicher Aufgaben in Mathematik erhöht wird. Nach den Annahmen der Autoren/-innen beeinflussen negative und positive Leistungsstereotypen den regulatorischen Fokus der Selbstregulation. Konsistent mit den Befunden der vorliegenden Studie scheint es so, dass ein *promotion focus* durch die *stereotype reactance* induziert werden kann. Die Entgegenwirkung gegen das negative Stereotyp der sozialen Gruppe von Jungen ähnelt sehr stark der Einnahme eines *promotion focus* (Annähe-

rungsstrategien, Risikobereitschaft). Jungen der Experimentalgruppe sind weniger gewillt, leichte Aufgaben auszuwählen, um zu brillieren (Annäherungs-Leistungsziel) oder um mögliche Defizite zu kaschieren (Vermeidungs-Leistungsziel), was einem *prevention focus* (Fehler- und Risikovermeidung, vgl. Evans & Petty, 2003; Crowe & Higgins, 1997) entsprechen würde. Ob die Auslösung auch auf anderem Wege, als durch positive Stereotype geschehen kann (in diesem Fall: ob eine Auslösung durch *stereotype reactance* möglich ist), kann nach dem derzeitigen Forschungsstand allerdings nicht abschließend beantwortet werden, wenngleich die Befunde der vorliegenden Studie für diese Möglichkeit sprechen. Um dies genauer feststellen zu können, sollten zukünftige Arbeiten zu dieser Fragestellung neben Selbstberichten auch nicht-reaktive Messmethoden anwenden (z.B. Vick et al., 2008), um eine differenziertere Betrachtungsweise der Wirkmechanismen zu ermöglichen.

Limitationen. Neben den neuen Befunden, welche durch die vorliegende Studie geliefert werden, um bisherige Erkenntnisse zu Wirkungsweise und Konsequenzen einer Stereotypenbedrohung zu ergänzen und auf die bisher nicht untersuchte Gruppe von Jungen im Kontext Schule anzuwenden, ist die Aussagekraft der beschriebenen Ergebnisse aus den nachfolgend spezifizierten Gründen eingeschränkt. Zunächst bleibt kritisch zu reflektieren, dass das zum Einsatz gebrachte Messinstrument von Leggett (1985) eine zwar hochgradig ökonomische, wenngleich aber auch eine stark vereinfachende Messung von motivationalen Zielorientierungen vornimmt. Es bleibt somit zu überprüfen, ob sich die gefundenen Ergebnisse unter Anwendung eines ausführlicheren Messinstruments, wie z.B. dem SELLMO-Fragebogen (Skalen zur Erfassung von Lern- und Leistungsmotivation; Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002) oder dem PALS (*Patterns of Adaptive Learning Strategies*; Midgley et al., 2000) replizieren lassen. Dies würde es ferner ermöglichen nicht nur Häufigkeitsanalysen durchzuführen, d.h. festzustellen, welche Untersuchungsgruppe welche motivationale Orientierung auswählt, sondern eine graduelle Abstufung der Ausprägung von Lern- und Leistungsmotivation vorzunehmen. Weiterhin könnte damit genauer geklärt werden in welchem Ausmaß die vorgefundenen Effekte vorliegen, d.h. eine Berechnung einer Effektstärke wäre möglich, ebenso wie die Anwendung weiterführender inferenzstatistischer Verfahren deren Voraussetzungen die vorliegende Untersuchung bereits das Mindest-Skalenniveau betreffend nicht erfüllen kann.

Weiterhin werden in beiden präsentierten Untersuchungen nur in einem geringen Ausmaß zusätzliche Leistungsdaten erhoben. Zwar wurde in einer Studie das initiale Leistungsniveau gemessen, jedoch fand keine Posttest-Messung statt. Eine so angelegte Messung würde es ermöglichen, die Effekte durch Stereotypenbedrohung genauer zu spezifizieren und in einen gemeinsamen Wirkzusammenhang zu bringen, da aus den vorliegenden Untersuchungsdaten nicht abgeleitet werden kann, inwiefern a) tatsächlich anspruchsvollere Aufgaben von Jungen der Experimentalbedingung in Mathematik bearbeitet werden, b) dies zu verbesserten Lernleis-

tungen führt und c) mit einem längerfristig gesteigerten Kompetenzerwerb in Mathematik einher geht.

Abschließend weisen die durch Smith (2004) und Thoman et al. (2013) postulierten Modelle darauf hin, dass jenseits der motivationalen Zielorientierungen weitere Bereiche beeinflussend wirken können (z.B. Identifikation mit der Leistungsdomäne; Smith & White, 2001) bzw. beeinflusst werden (z.B. Selbstwert; Cohen & Garcia, 2005 / Bereitschaft zu *self-handicapping*; Keller, 2002). Um die möglicherweise komplexen Wirkrichtungen angemessener prüfen zu können wäre ein vermehrter Einbezug ebendieser Kontextfaktoren wünschenswert. Dies könnte dazu beitragen die Auswirkungen einer Stereotypenbedrohung in ihrer Vielschichtigkeit besser verstehen zu können, um darauf aufbauend interventionelle Maßnahmen zu erarbeiten und zu erproben, welche nicht nur punktuell eine lernförderlichere Motivation etablieren, sondern umfassender auch weitere lernrelevante Facetten und Wohlbefinden in Lehr-Lern-Settings positiv beeinflussen können. Um schlussendlich auch die Frage beantworten zu können, ob durch die Beeinflussung motivationaler Zielorientierungen infolge einer Stereotypenbedrohung eine gewisse Dauerhaftigkeit in den negativen Konsequenzen auf der Leistungsebene entsteht, wäre zur Absicherung dieser derzeit noch offenen Forschungsfrage zusätzlich eine Längsschnittuntersuchung erforderlich.

8. Studie 5: Wiederherstellung positiver Distinktheit I

Aufwertung von Männlichkeit und Abwertung von Weiblichkeit als Bewältigungsstrategie auf das Stereotyp des männlichen Schulversagers

8.1. Überblick und Fragestellung

In den vorausgehenden Untersuchungen dieser Arbeit konnte aufgezeigt werden, dass die Konfrontation von männlichen Schülern mit einem Stereotyp, welches ihr schulisches Versagen propagiert, die Leistungsfähigkeit und die Motivation zur Bearbeitung von entsprechenden Aufgaben differenziell beeinflussen und hemmen kann (*Studien 2, 3 & 4*, vgl. auch Hartley & Sutton, 2013). Aus der im Kapitel 4 ausführlicher vorgestellten Theorie der sozialen Identität (Tajfel & Turner, 1986) ist abzuleiten, dass eine ebensolche Bedrohungssituation nicht nur Einfluss auf die Qualität und Motivation einer Aufgabenbearbeitung haben, sondern vielmehr auch zu einer Beeinflussung des sozialen Gefüges zwischen den beteiligten Gruppen führen kann. Genauer wird postuliert, dass die Wahrnehmung einer negativen Diskrepanz zu einer Außengruppe, welche durch die negative Darstellung von Jungen in der Schule in den vorliegenden Untersuchungen unterstützt wird, zu einer Tendenz führt, die Überlegenheit der eigenen Gruppe wieder stärker herauszustellen (Wiederherstellung positiver Distinktheit). Die Mechanismen, welche die Theorie zur Absicherung einer positiven Abgrenzung der Eigengruppe vorschlägt, sind bisher in der Hauptsache im Erwachsenenalter und vermehrt im organisationalen Kontext angewendet worden (Hogg & Terry, 2001; van Knippenberg, 2000; van Knippenberg, van Dick & Tavares, 2007). Im Rahmen der folgenden Studie soll eine durch die *Social Identity Theory* (Tajfel & Turner, 1986) vorgeschlagene Strategie zur Wiederherstellung positiver Distinktheit auf den Kontext Jungen in der Schule angewendet werden: Übertragen auf die Problemstellung dieser Arbeit ist anzunehmen, dass die Wahrnehmung von sich wandelnden Intergruppenbeziehungen zum Nachteil der Gruppe der Jungen zu einer Entstehung von ungesicherten sozialen Identitäten in der Schule führt. Hieraus erwächst das Bedürfnis, den vorher geltenden Status der Überlegenheit der Eigengruppe wiederherzustellen, was sozialen Wettbewerb wahrscheinlich macht (Jackson & Smith, 1999). In der Tat kann unter Bezugnahme auf die statistischen Daten zur Bildungsbeteiligung und zum Bildungserfolg gefunden werden, dass sich hier in den vergangenen Jahren ein deutlicher Wandel der Intergruppenbeziehungen bestätigen lässt. Während in früheren Kohorten mehr Männer höhere Schulformen besuchten und höherwertige Abschlusszertifikate erreichten, besuchen aktuell mehr Mädchen höhere Schulformen einhergehend mit einem Erwerb höher qualifizierender Abschlüsse (Statistisches Bundesamt, 2014a). Unter der Maßgabe, dass dieser Wandel bekannt ist oder aber im Verlauf der Untersuchung bewusst gemacht wird, wird es somit wahrscheinlich, dass die als benachteiligt dargestellte Gruppe der Jungen Strategien zur Wiederherstellung von positiver Distinktheit anwendet. Eine erste im Rahmen dieser Studie zu prüfende Strategie aus der Theorie stellen Aufwertungsprozesse für die Eigengruppe vs. Abwertungsprozesse für die Außengruppe dar. Die Aufwertung der Eigengruppe wird

hier angenommen als eine verstärkt positive Einstellung gegenüber Maskulinität und eine Honorierung männlichen Verhaltens, insbesondere im schulischen Kontext. Zur Abwertung der Außengruppe wird vermutet, dass hierbei weibliches Verhalten und Femininität generell abgelehnt und derartiges Verhalten bzw. eine derartige Gruppenzugehörigkeit sanktioniert wird.

Das Konzept des ambivalenten Sexismus (Glick & Fiske, 1996) bietet eine Möglichkeit, die Abwertung von Weiblichkeit differenziert auf zwei Faktoren zu betrachten. Die Autoren/-innen der Theorie gehen davon aus, dass Sexismus nicht ausschließlich durch hostile Aspekte ggü. Weiblichkeit charakterisiert werden kann, sondern dass dieser auch aus vermeintlich positiven Überzeugungen ggü. Frauen entstehen kann. Entsprechend dieser Grundannahmen werden zwei Komponenten postuliert (und empirisch nachgewiesen), aus welchen sich Sexismus definiert: Einerseits drückt sich Sexismus durch eine hostile-feindselige Komponente aus, welche die höhere strukturelle Macht von Männern betont und daraus eine dominante Stellung des Mannes im Geschlechterverhältnis ableitet (vgl. Bierhoff, 2006). Inhaltlich schlagen sich derartige Einstellungen in Überzeugungen nieder, dass Frauen auf männliche Hilfe und Unterstützung im Leben angewiesen sind und ohne diese nicht in der Lage sind, allen Anforderungen des alltäglichen Lebens gerecht zu werden. Eine zweite, mit positiven Gefühlen verbundene Komponente drückt eine benevolent-wohlwollende Einstellung gegenüber Frauen aus und basiert auf der höheren dyadischen Macht von Frauen, da Männer u.a. zur Reproduktion auf sie angewiesen sind. Dies spiegelt sich inhaltlich z.B. darin wider, dass die Notwendigkeit des Schutzes von Frauen durch Männer, auf heterosexuelle Sexualität bezogene Intimität und die gegenseitige Ergänzung von Mann und Frau in familialen Rollen betont werden (Glick & Fiske, 1996). Die soeben beschriebenen Hauptkomponenten lassen sich Glick & Fiske (1996) folgend zusätzlich in je drei Subdimensionen aufgliedern, nämlich: Paternalismus, also das Machtgefüge zwischen Männern und Frauen (hostil: dominant, i.S. einer Infragestellung, dass Frauen ohne männliche Unterstützung vollständig kompetente Individuen sind / benevolent: protektiv, d.h. die Betonung, dass Frauen männlichen Schutz benötigen), Geschlechterdifferenzierung, also physische und psychologische Merkmale zur genauen Abgrenzung beider Geschlechter (hostil: kompetitiv, also Betonung der Richtigkeit höherer struktureller und politischer Macht von Männern / benevolent: komplementär, d.h. Fokussierung darauf, dass Frauen andere Eigenschaften als Männer haben, und dass sich die Eigenschaften von Männern und Frauen sinnvoll ergänzen) und Heterosexualität, welche die Basis einer romantischen Liebesbeziehung zwischen Mann und Frau bildet (hostil: aggressiv, i.S. sexueller Dominanz / benevolent: intim, d.h. Betonung, dass die Partnerschaft mit einer Person des anderen Geschlechts Grundlage eines erfüllten Lebens ist) (vgl. auch Rau, 2013). Insgesamt bietet der Ansatz von Glick & Fiske (1996, des Weiteren: 1999, 2001) eine systemische Grundlage zur Erklärung von Entstehung, Aufrechterhaltung und der gesellschaftlichen, systemstabilisierenden Konsequenzen von Sexismus. Für die vorliegende Arbeit scheint

das Konzept des ambivalenten Sexismus besonders geeignet, um zu überprüfen, ob eine Bedrohung der bisher geltenden Intergruppenrelationen in einer (hostilen und/oder benevolenten) Entwertung von Weiblichkeit resultiert. Darüber hinaus erlaubt diese Art der Erfassung von Sexismus im Rahmen einer experimentellen Manipulation festzustellen, ob Sexismus als psychologisches Konstrukt (im Sinne eines „state“) von situationalen Merkmalen beeinflusst werden kann.

Folgende Hypothesen sollen im Rahmen der vorliegenden Studie überprüft werden:

- H9* Jungen der Experimentalgruppe bewerten relativ zu Jungen der Kontrollgruppe eine stark männliche Personenbeschreibungen (Vignette) positiver.
- H10* Jungen der Experimentalgruppe bewerten relativ zu Jungen der Kontrollgruppe eine wenig männliche Vignette negativer.
- H11* Jungen der Experimentalgruppe reagieren auf das Treatment relativ zu Jungen der Kontrollgruppe mit stärkerer hostile-sexistischer (Weiblichkeit abwertenden) Einstellungen.
- H12* Mädchen weisen zwar im Vergleich zu Jungen generell weniger hostile-sexistische Einstellungen auf, divergieren aber in ihren hostile-sexistischen Einstellungen nicht in Abhängigkeit von der Untersuchungsbedingung.

Zur Komponente des benevolenten Sexismus im Jugendalter, der stark vom Grad der Ausbildung intimer Paarbeziehungen abhängig ist (de Lemus, Moya & Glick, 2010), soll auf die Ableitung konkreter Hypothesen verzichtet und die resultierenden Ergebnisse explorativ analysiert werden. So gehen auch Glick & Hilt (2000) davon aus, dass aufgrund der Tatsache dass heterosexuelle, intime Beziehungen für Kinder noch keine Rolle spielen lediglich die feindselige Komponente des Sexismus ausgeprägt ist. Dies wird z.B. sichtbar in selbstgewählter Geschlechtertrennung in kindlichen Spiel- und Lerngruppen, wobei aus den unterschiedlichen Interaktionsstilen von Mädchen und Jungen ein Machtgefälle entsteht, aus dem Glick und Hilt (2000) zufolge feindselig sexistische Einstellungen entstehen (dominanter Paternalismus bei Jungen, Verübeln von Paternalismus bei Mädchen).

Zusammenfassend soll eine Strategie der Theorie der sozialen Identität (Tajfel & Turner, 1986) auf das vergleichsweise neue Problem der „Jungen als Bildungsverlierer“ angewendet werden, um zu überprüfen ob Jungen nach einer Bedrohungssituation, die einen Wandel der geltenden Intergruppenbeziehungen verdeutlicht, Maskulinität auf- und Femininität abwerten, um ihre bedrohten sozialen Identitäten abzusichern.

8.2. Methode

Vortests zur Vignettenerstellung. Zur Erstellung von Vignetten, auf denen jeweils ein stark männlichkeits-inszenierender Junge sowie ein Junge, welcher sich durch vornehmlich geschlechtsneutrale Verhaltensweisen auszeichnet abgebildet wird, wurden verschiedene Verhaltensweisen (entnommen von Heyder & Kessels, 2012) in einem Vortest dahingehend überprüft, inwiefern diese Verhaltensweisen a) zur Inszenierung von Männlichkeit geeignet sind und b) wie sozial erwünscht das abgefragte Verhalten wahrgenommen wird. Hierzu wurden insgesamt 98 Verhaltensweisen vorgelegt (davon 37 männlichkeitsinszenierend, z.B. „sehr breitbeinig sitzen“; 39 weiblichkeitsinszenierende Eigenschaften, z.B. „sich auf der Schultoilette schminken“ und 22 neutrale Verhaltensweisen, z.B. „im Unterricht Kaugummi kauen“). Zur Erfassung des Inszenierungspotential wurden 101 Schüler/-innen (57 Jungen, 44 Mädchen, Alter $M = 14.99$ Jahre, $SD = .82$, range 13-17 Jahre) der Klassenstufe 9 einer integrierten Sekundarschule in einem Außenbezirk der Stadt Berlin befragt. Das Antwortformat reichte von 1 (entspricht stark zur Inszenierung von Männlichkeit geeignet) über 4 (geschlechterneutrale Verhaltensweise) bis 7 (stark zur Inszenierung von Weiblichkeit geeignet). Zur Erfassung der sozialen Erwünschtheit wurden diese Daten um 72 Schüler/-innen (35 Jungen, 36 Mädchen, 1 fehlender Wert, Alter $M = 14.82$. $SD = 0.57$, range 14-16 Jahre) der Klassenstufe 9 eines Berliner Gymnasiums ergänzt, welche dieselben Verhaltensweisen hinsichtlich deren sozialen Erwünschtheit bewerteten (1 = außerordentlich erwünscht, 4 = weder/noch, 7 = vollkommen unerwünscht).

Im Ergebnis wurde aus den vorgetesteten Verhaltensweisen diejenigen für die weitere Befragung ausgewählt, die einerseits am stärksten zur Inszenierung von Männlichkeit (Mittelwert nahe 1) oder aber am geschlechtsneutralsten (Mittelwert nahe 4) bewertet wurden und eine vergleichbare soziale Erwünschtheit aufweisen um zu vermeiden, dass die starke Eignung hinsichtlich des Inszenierungspotentials durch Unterschiede in der sozialen Erwünschtheit von bestimmten Verhaltensweisen konterkariert wird. Konkret wurden aus den Bewertungen beider Vorstudien folgende Verhaltensweisen für die Verwendung in der Hauptstudie ausgewählt:

- a) Männlichkeitsinszenierender Junge: Freunde aufwändig rituell begrüßen, z.B. „Ghettofaust“, in der Pause Armdrücken machen, bewusst breitbeinig dasitzen, mit sportlichen Erfolgen prahlen, andere beim Sport übertreffen wollen, Schmerzen ertragen und bewusst herunterspielen
- b) Junge mit neutralen Verhaltensweisen: andere mit eigenen Redebeiträgen übertrumpfen wollen, in der Pause Musik mit dem Handy hören, beim Melden mit den Fingern schnippen, andere Mitschüler/-innen zurechtweisen, niemals Stifte und Arbeitsmaterial verleihen.

Die so extrahierten Verhaltensweisen wurden in eine kurze Geschichte eines neuen Mitschülers eingebunden, der neu in die Klassengemeinschaft gekommen ist und die o.g. Verhaltensweisen häufig aufzeigt.

Teilnehmende. Für die Hauptuntersuchung konnten insgesamt 88 Schüler/-innen der Klassenstufe 9 einer integrierten Gesamtschule gewonnen werden. Insgesamt 8 Fragebögen waren nicht oder nur unzureichend bearbeitet (>75% fehlende Werte) und wurden daher aus den weiteren Analysen ausgeschlossen. Das Alter lag bei $M = 14.60$ Jahren, $SD = 0.73$, $range = 14-17$ Jahre). Insgesamt entfielen von 38 teilnehmenden Jungen 22 auf die Experimentalbedingung und 16 auf die Kontrollbedingung. Von den 42 teilnehmenden Mädchen wurden 19 der Experimental- und 23 der Kontrollgruppe zugewiesen.

Hinsichtlich des Migrationshintergrundes gaben drei von 35 männlichen (3 fehlende Werte) sowie fünf von 41 weiblichen (1 fehlender Wert) Teilnehmenden an, im Ausland geboren zu sein (direkter Migrationshintergrund). Wird, wie in den vorhergehenden Studien, ergänzend die Sprache mit den Eltern als Indikator für einen (indirekten) Migrationshintergrund herangezogen zeigt sich ein leicht höherer Anteil von Migranten/-innen: Von den Personen, die hierzu Angaben gemacht haben, sprachen 10 von 37 Jungen (1 fehlender Wert) und 13 von 42 Mädchen mit ihren Eltern eine andere Sprache als Deutsch. Beide Indikatoren unterscheiden sich nicht signifikant zwischen den Geschlechtern, alle $\chi^2 < 0.26$, $p > .61$. Betrachtet nach Untersuchungsbedingung entfielen auf die Experimentalgruppe eine Person mit direkter Migrationserfahrung von insgesamt 38 Angaben (3 fehlende Werte), in der Kontrollgruppe gaben sieben von 38 Personen (bei einem fehlendem Wert) an, im Ausland geboren zu sein. Mit einem $\chi^2 (1, N = 76) = 5.03$, $p = .03$ wird dieser Unterschied in Richtung eines höheren Anteils von Personen mit direkter Migrationserfahrung in der Kontrollgruppe statistisch bedeutsam. Unter Berücksichtigung der mit den Eltern gesprochenen Sprache gaben 12 von 40 Personen (1 fehlender Wert) in der Experimentalgruppe und 11 von 39 Personen in der Kontrollgruppe an, mit den Eltern eine andere Sprache als Deutsch zu sprechen, mit $\chi^2 < 0.04$, $p = .86$ indiziert dieser Indikator für (indirekte) Migrationserfahrung somit keinen Unterschied zwischen den Untersuchungsgruppen mehr.

Der sozioökonomische Hintergrund lag mit $M = 4.95$ ($SD = 1.26$, $range = 2-7$) von 7 verfügbaren Gütern auf einem durchschnittlichen Niveau, wobei zwischen Jungen und Mädchen mit $M_{\text{Jungen}} = 5.03$, $SD = 1.28$; $M_{\text{Mädchen}} = 4.88$, $SD = 1.25$ kein bedeutsamer Unterschied auf den sozioökonomischen Indikatorvariablen existiert ($F < 1$, n.s.). Der Vergleich zwischen den Untersuchungsbedingungen deutet jedoch darauf hin, dass mit $M_{\text{EG}} = 5.23$, $SD = 1.23$ und $M_{\text{KG}} = 4.67$, $SD = 1.24$, $F(1,75) = 4.01$, $p < .05$, $\eta^2 = .05$ in der Experimentalgruppe von einer höheren Verfügbarkeit der abgefragten Güter auszugehen ist. Somit bleibt abschließend festzuhalten, dass im Hinblick

auf die Ergebnisse der initialen Analysen der Teilnehmenden ein Einfluss der Ungleichverteilung in den Untersuchungsbedingungen von Personen mit direkter Migrationserfahrung sowie hinsichtlich des sozioökonomischen Status nicht auszuschließen ist. Daher erscheint es sinnvoll, die Ergebnisanalyse ggf. um korrektive Analyseverfahren (z.B. Kovarianzanalyse, *ANCOVA*) zu ergänzen.

Prozedur und Material. Mittels einer standardisierten Instruktion wurden die Teilnehmenden begrüßt und über den Ablauf der Untersuchung informiert. Nachdem sich alle Teilnehmer/-innen unter Zusicherung vollständiger Anonymität freiwillig zur Teilnahme an der Untersuchung bereit erklärt haben wurden die vorbereiteten Testhefte in randomisierter Abfolge verteilt. Eine weitere, kurze Instruktion machte die Schüler/-innen mit der Anwendung der Testhefte und vor allem mit den darin enthaltenen Aufgabenformaten sowie den psychometrischen Skalen vertraut. Nachdem alle offenen Fragen vollständig geklärt wurden und eine Entlohnung von 10 Euro für die Klassenkasse je ernsthaft bearbeiteten Fragebogen zur Erhöhung der *compliance* in Aussicht gestellt wurde, begann die eigentlich Arbeit mit dem Testheft. Zunächst wurde eine Übungsaufgabe präsentiert, in welcher die Schüler/-innen eine kurze Statistik zum Mediengebrauch von Jugendlichen in Deutschland interpretieren und damit einhergehend Daten aus einer Abbildung auswerten und in entsprechende Schlussfolgerungen integrieren sollten. Diese Aufgabe diente der Einübung notwendiger Arbeitsschritte und Vorgehensweisen für die Auswertung der folgenden Treatment-Aufgabe, wurde den Schüler/-innen allerdings maskiert als Leistungsaufgabe präsentiert. Nach Beantwortung von vier Kontrollfragen zur Aufgabe wurde das Treatment präsentiert. In der Experimentalgruppe war eine mit der Übungsaufgabe vergleichbare Statistik zu bearbeiten, aus der hervorging, dass a) Jungen wesentlich seltener das Gymnasium besuchen als Mädchen; b) Jungen deutlich schlechtere Durchschnittsnoten erwerben als Mädchen und c) Jungen heute – im Unterschied zu früheren Kohorten – eher schulisch geringer wertige Abschlusszertifikate erhalten als Mädchen. Die Kontrollgruppe bearbeitete eine Aufgabe vergleichbarer Länge und Schwierigkeit ohne einen Bezug auf Geschlechterunterschiede im Bildungswesen (Wandel von Schulbesuchszahlen in Hamburg und Bayern). Wie in der Übungsaufgabe trainiert folgten auch auf die Treatment-Aufgabe drei Kontrollfragen, um sicherzustellen, dass alle notwendigen Informationen aus dem dargebotenen Material entnommen und folgerichtig interpretiert wurden. Darauf aufbauend wurden in randomisierter Reihenfolge die zwei auf den Daten des Vortests basierenden Vignetten präsentiert: einmal ein Junge, welcher sich im schulischen Kontext stark männlich inszeniert sowie eine zweite Vignette eines Jungen, der sich im schulischen Kontext geschlechtsneutral inszeniert. Diese Vignetten waren jeweils hinsichtlich Wärme (z.B. tolerant, gutherzig, aufrichtig) und Kompetenz (z.B. intelligent, selbstsicher, unabhängig) auf dem Messinstrument von Fiske, Cuddy, Glick & Xu (2002) auf fünfstufigen semantischen Differentialen zu bewerten, zusätzlich wurde mit zwei zusätzlichen Fragen erho-

ben, wie maskulin vs. feminin die beschriebene Person auf der Vignette eingeschätzt wird. Die Skala von Fiske et al. (2002) basiert auf der Grundannahme, dass die zwei Dimensionen wahrgenommene Wärme und Kompetenz zwei primäre Dimensionen sind, auf welchen sich Personen beschreiben und Stereotype abbilden lassen, wobei häufig gemischte Cluster auftreten (hohe Kompetenz, aber geringer Wärme und v.v.). Diese Einschätzungen ziehen wiederum verschiedene emotionale Konsequenzen nach sich (z.B. Mitleid bei wahrgenommener Inkompetenz). In zahlreichen Studien mit verschiedenen Stichproben und sozialen Gruppen konnte nachgewiesen werden, dass die Wahrnehmung von Fremdgruppen den Dimensionen Wärme und Kompetenz zugeordnet werden konnte und diese häufig gegensätzliche Ausprägungen aufweisen, was dieses Instrument als besonders geeignet für die ökonomische Erfassung von (stereotypen) Personenbewertungen erscheinen lässt.

Es folgten vier weitere Kontrollfragen zum Treatment mit dem Ziel, dessen Wirkung durch eine erneute Auseinandersetzung mit den vorgelegten Materialien aufzufrischen. Hiernach wurde das Ambivalente Sexismus-Inventar für Jugendliche (Rau & Hannover, 2008) zur Erfassung sexistischer Einstellungen vorgelegt (Skalierung 1 = stimme gar nicht zu, 5 = stimme voll und ganz zu). Nach der Erfassung demografischer Daten (wie z.B. Alter, Geschlecht, Herkunft und sozioökonomischer Hintergrund) erfolgte ein sorgfältiges *Debriefing* mit vollständiger Aufklärung über Ziel und Inhalt der Untersuchung sowie der Möglichkeit, offen im Klassenraum oder im Einzelgespräch etwaige Fragen oder Probleme zu klären. Keine der teilnehmenden Personen berichtete nach Abschluss der Untersuchung einen weiteren Gesprächs- oder Klärungsbedarf. Die Untersuchung hatte einschließlich Instruktion und Nachbereitung sowie der symbolischen Übergabe der Entlohnung eine Gesamtdauer von ca. 75 Minuten.

8.3. Ergebnisse

Initiale Analysen. In einem ersten Schritt wurde erneut sichergestellt, dass die in die Analyse aufgenommenen Probanden/-innen die Treatment-Materialien sorgfältig bearbeitet haben und damit einhergehend die Fragen zur Manipulationsüberprüfung in einem hinreichenden Ausmaß (mindestens 50% korrekt) beantwortet haben. Eine ANOVA mit Messwiederholung (t1: vor Bearbeitung der Vignetten, t2: vor Bearbeitung des ambivalenten Sexismus-Inventars für Jugendliche) zum Anteil korrekter Antworten ergibt mit $F(1,76) = 3.84, p = .05, \eta^2 = .05$ einen signifikanten Haupteffekt für den Messzeitpunkt in Richtung einer höheren Anzahl korrekt beantworteter Fragen zur Manipulationsüberprüfung zum zweiten Messzeitpunkt ($M = 65.94\%$, $SD = 21.88\%$) im Vergleich zum ersten Messzeitpunkt ($M = 60.00\%$, $SD = 12.95\%$). Weiterhin indiziert ein signifikanter Interaktionseffekt, dass zum zweiten Messzeitpunkt die Probanden/-innen der Experimentalgruppe mit $M = 73.17\%$, $SD = 22.63\%$ einen substantiell höheren

Anteil korrekt beantworteter Fragen aufweisen, als Teilnehmende in der Kontrollgruppe ($M = 58.33\%$, $SD = 18.44\%$) oder Teilnehmende beider Untersuchungsgruppen zum ersten Messzeitpunkt ($M_{EG} = 60.98\%$, $SD = 13.75\%$, $M_{KG} = 58.97\%$, $SD = 12.15\%$), $F(1,76) = 6.39$, $p = .01$, $\eta^2 = .08$. Zusätzlich wird die Dreifachinteraktion mit den Faktoren Messzeitpunkt, Bedingung und Geschlecht marginal signifikant, $F(1,76) = 3.60$, $p = .06$, $\eta^2 = .05$ und deutet darauf hin, dass Jungen der Experimentalgruppe zum zweiten Messzeitpunkt mehr Kontrollfragen richtig beantwortet haben ($M = 77.27\%$, $SD = 21.70\%$) als alle sonstigen Probanden/-innen zum ersten und zum zweiten Messzeitpunkt. Zur weiteren Klärung dieser Befunde wurden post hoc T-Tests durchgeführt. Die ersten Fragen zur Manipulationsüberprüfung (vor der Einschätzung der Vignetten), wiesen mit einem mittleren, prozentualen Anteil korrekter Antworten von $M_{t1} = 60.00\%$ ($SD = 12.95\%$, $range = 50-100\%$) eine ausreichende Anzahl richtiger Antworten auf. Bei den zweiten Fragen zum *manipulation check*, welche vor der Bearbeitung des ambivalenten Sexismus-Inventars für Jugendliche abgefragt wurden, fiel der Anteil richtiger Antworten mit $M_{t2} = 65.94\%$ ($SD = 21.88\%$, $range = 50-100\%$) signifikant höher und damit auf einem zufriedenstellenden Niveau aus, $\Delta M = 5.94$, $t(79) = -2.19$, $p < .05$. Weitere Einzelvergleiche beider Geschlechter hinsichtlich des Anteils korrekter Antworten indizieren bei beiden Messzeitpunkten zur Manipulationsüberprüfung keine bedeutsamen Unterschiede zwischen Jungen und Mädchen, $M_{Jungen_{t1}} = 60.53\%$ ($SD = 13.79\%$), $M_{Mädchen_{t1}} = 59.52\%$ ($SD = 12.29\%$), $M_{Jungen_{t2}} = 67.76\%$ ($SD = 21.68\%$), $M_{Mädchen_{t2}} = 64.29\%$ ($SD = 22.18\%$); $t(78) < .71$, $p > .48$. Der Einzelvergleich zum ersten Messzeitpunkt nach Untersuchungsbedingung weist mit $M_{EG_{t1}} = 60.98\%$ ($SD = 13.75\%$) sowie $M_{KG_{t1}} = 58.97\%$ ($SD = 12.15\%$), $t(78) = .67$, $p = .49$ keine bedeutsamen Unterschiede auf, wohingegen zum zweiten Messzeitpunkt (vor Bearbeitung des ASI) deutliche Differenzen in Richtung eines höheren Anteils korrekter Antworten in der Experimentalgruppe auftreten: $M_{EG_{t2}} = 73.17\%$ ($SD = 22.63\%$), $M_{KG_{t2}} = 58.33\%$ ($SD = 18.44\%$), $t(78) = 3.22$, $p < .01$, $d = 0.71$, wobei insbesondere Jungen der Experimentalgruppe zum zweiten Messzeitpunkt mit $M = 77.27\%$, $SD = 21.70\%$ deutlich mehr Kontrollfragen korrekt beantworteten als Jungen der Kontrollgruppe ($M = 54.69\%$, $SD = 13.60\%$), $t(36) = 3.67$, $p < .001$, $d = 0.99$. Diese Tatsache ist im Kontext der vorliegenden Untersuchung dahingehend vergleichsweise unproblematisch, als dass es in der Experimentalgruppe und insbesondere bei Jungen, wo ein höherer Anteil korrekter Antworten vorliegt, von zentraler Bedeutung für das Gelingen der Untersuchung ist, dass die Probanden/-innen das vorgelegte Material mit höchster Sorgfalt verarbeiten, was gelungen zu sein scheint. Ferner konnte in der Kontrollgruppe nichtsdestotrotz auch ein ausreichend hoher Anteil korrekter Antworten erreicht werden.

Des Weiteren wurde im Rahmen der Beurteilung der Vignetten überprüft, inwiefern diese tatsächlich als maskulin vs. feminin wahrgenommen werden. Der paarweise Vergleich beider Vignetten zeigt mit einer Mittelwertsdifferenz von $\Delta M = 0.84$, $t(76) = 6.75$, $p < .001$, dass die erste Vignette mit einer durchschnittlichen Bewertung von $M = 3.68$ ($SD = 1.12$) signifikant maskuli-

ner beurteilt wurde als die zweite, eher geschlechterneutral konzipierte Vignette mit $M = 2.84$ ($SD = .53$).

Analysen zur Vignetten-Bewertung. Basierend auf den Analysen zur Manipulationsüberprüfung und zur Absicherung der konstruierten Vignetten kann somit angenommen werden, dass eine angemessene Grundlage für die weiterführenden Analysen geschaffen wurde. Um nunmehr feststellen zu können, ob sich getriggert durch das experimentelle Treatment die Bewertungen der Vignetten eines stark maskulinen vs. geschlechtsneutral dargestellten Jungen systematisch voneinander unterschieden, wurden aus den entsprechenden Einzelfragen die Skalen Wärme (Vignette 1: $\alpha = .72$; Vignette 2: $\alpha = .78$) und Kompetenz (Vignette 1: $\alpha = .80$; Vignette 2: $\alpha = .81$) gebildet und je Skala einer 2 (Bedingung) \times 2 (Geschlecht) \times (2) (Messwiederholungsfaktor: männliche vs. wenig männliche Vignette) *mixed model*-ANOVA unterzogen. Auf der Skala Wärme/emotionale Nähe zeigen die Ergebnisse der Varianzanalyse einen signifikanten Haupteffekt auf dem Messwiederholungsfaktor, $F(1,74) = 17.62$, $p < .001$, $\eta^2 = .19$, in Richtung einer generellen, unabhängig von Experimentalbedingung und Geschlecht, besseren Bewertung der ersten, stark männlichkeitsinszenierenden Vignette, $M_{V1} = 2.89$ ($SD = .81$), $M_{V2} = 2.40$ ($SD = 1.00$). Zusätzlich erreicht der Haupteffekt für das Geschlecht der Probanden/-innen ein marginales Signifikanzniveau in Richtung einer besseren Bewertung der Vignetten durch Jungen, $M = 2.80$ ($SD = .84$), als durch Mädchen, $M = 2.49$ ($SD = .92$), $F(1,74) = 3.30$, $p = .07$, $\eta^2 = .04$. Weder die Interaktionsterme aus Messzeitpunkt und Geschlecht oder Experimentalbedingung, noch die Dreifachinteraktion erreichen das notwendige Signifikanzkriterium, $F(1,74) < 1.42$, $p > .24$. Post hoc-Testungen sichern diesen Befund ab und indizieren keine signifikanten Unterschiede auf beiden Vignetten zwischen Jungen der Experimental- und Kontrollgruppe ($M_{EG_V1} = 2.94$, $SD = .83$, $M_{KG_V1} = 3.10$, $SD = 1.01$, $M_{EG_V2} = 2.69$, $SD = .89$, $M_{KG_V2} = 2.47$, $SD = .62$, $t < 1$, $p > .40$) oder zwischen Jungen und Mädchen der Experimentalgruppe ($M_{Jungen_V1} = 2.94$, $SD = .83$, $M_{Mädchen_V1} = 2.78$, $SD = .65$, $M_{Jungen_V2} = 2.69$, $SD = .89$, $M_{Mädchen_V2} = 2.12$, $SD = 1.43$, $t(39) < 1.54$, $p > .13$). Um die initialen Unterschiede bezüglich des direkten Migrationshintergrundes und des sozioökonomischen Status stärker zu berücksichtigen, wurde ergänzend eine Kovarianzanalyse berechnet, um beide Einflussfaktoren statistisch besser zu kontrollieren. Im Ergebnis dieses Analyseschrittes lässt sich lediglich der tendenziell bedeutsame Haupteffekt auf dem Faktor Geschlecht in gleicher Richtung und Stärke replizieren, $F(1,69) = 3.10$, $p = .08$, $\eta^2 = .04$. Darüber hinaus zeigen sich weder signifikante Haupteffekte, noch statistisch bedeutsame Interaktionen, alle $F < 1$, n.s., was darauf hinweist, dass auch unter Kontrolle der initial ungleich verteilten Variablen sozioökonomischer Status und ethnische Herkunft nur eine geringfügige Änderung des Befundmusters festzustellen ist.

Auf der Skala wahrgenommene Kompetenz der auf den Vignetten beschriebenen Personen konnten im Rahmen der Varianzanalyse mit Messwiederholung keine statistisch bedeutsamen Haupt- oder Interaktionseffekte festgestellt werden, $F(1,74) < 1.68, p > .20$. Die erste, männlichkeitsinszenierende Vignette bewerteten Jungen der Experimentalgruppe ($M = 3.09, SD = .86$) vergleichsweise ebenso kompetent wie Jungen der Kontrollgruppe ($M = 2.77, SD = .77$) oder Mädchen der Experimental- ($M = 2.92, SD = .70$) oder Kontrollgruppe ($M = 2.88, SD = .63$). Außerdem divergieren die Kompetenzeinschätzungen auf der zweiten, geschlechterneutraleren Vignette nicht bedeutsam zwischen den Geschlechtern und Untersuchungsbedingungen ($M_{\text{Jungen_EG}} = 2.90, SD = .84, M_{\text{Jungen_KG}} = 2.81, SD = .80, M_{\text{Mädchen_EG}} = 2.84, SD = .71, M_{\text{Mädchen_KG}} = 2.53, SD = .72$). Ergänzende post hoc-Einzelvergleiche unterstützen die varianzanalytischen Ergebnisse, dass weder zwischen Jungen der Experimental- und Kontrollgruppe, noch zwischen Jungen und Mädchen der Experimentalgruppe bedeutsame Unterschiede festzustellen sind, $t < 1.29, p > .21$. Erneut wurde eine Kovarianzanalyse zur Konstanthaltung der im Rahmen der initialen Analysen festgestellten Unterschiede zwischen den Geschlechtern und Untersuchungsbedingungen durchgeführt, welche allerdings ebenfalls keine signifikanten Haupt- oder Interaktionseffekte ergab, $F < 1.78, p > .19$ und damit die Ergebnisse der vorangegangenen Varianzanalyse bestätigt.

Ergebnisse zum Sexismus. Das nach Auffrischung des Treatments durch die Schüler/-innen bearbeitete ambivalente Sexismus-Inventar für Jugendliche (Rau & Hannover, 2008) besteht aus zwei Subskalen (vgl. S. 143): Einerseits dem hostilem, also feindseligen Sexismus ($\alpha = .89$), andererseits dem benevolentem, wohlwollenden Sexismus ($\alpha = .86$). Beide Komponenten des ambivalenten Sexismus wurden auf einem Messwiederholungsfaktor zusammengefasst und mit den Faktoren Experimentalbedingung und Geschlecht einer 2 (Bedingung) \times 2 (Geschlecht) \times (2) (Messwiederholungsfaktor: hostiler vs. benevolenter Sexismus) *mixed model*-ANOVA unterzogen. Die Ergebnisse verdeutlichen einen signifikanten Haupteffekt für den Messwiederholungsfaktor. Der Subskala hostiler Sexismus wird mit $M = 2.79$ ($SD = .91$) weniger zugestimmt als der Subskala benevolenter Sexismus, $M = 3.37$ ($SD = .78$), unabhängig von Geschlecht und Untersuchungsbedingung, $F(1,76) = 35.12, p < .001, \eta^2 = .32$. Ebenso weist eine signifikante Interaktion aus dem Messwiederholungsfaktor und dem Faktor Geschlecht, $F(1,76) = 39.85, p < .001, \eta^2 = .34$ darauf hin, dass Jungen der Komponente des hostilem Sexismus stärker zustimmen ($M = 3.47, SD = .58$) als Mädchen ($M = 2.18, SD = .69$), $t(78) = 8.99, p < .001, d = 2.02$. Auf der Komponente des benevolenten Sexismus ist mit $M_{\text{Jungen}} = 3.38, SD = .56$ und $M_{\text{Mädchen}} = 3.36, SD = .94$ kein bedeutsamer Unterschied festzustellen.

Diese Ergebnisse werden durch die signifikante Dreifachinteraktion ergänzt, $F(1,76) = 6.33, p < .05, \eta^2 = .08$: Jungen in der Experimentalgruppe stimmen der Komponente des hostilem

Sexismus stärker zu als Jungen der Kontrollgruppe ($M_{\text{Jungen_EG}} = 3.74, SD = .51, M_{\text{Jungen_KG}} = 3.09, SD = .47$, post hoc: $t(36) = 4.04, p < .001, d = 1.34$) und Mädchen in beiden Untersuchungsgruppen ($M_{\text{Mädchen_EG}} = 2.36, SD = .63, M_{\text{Mädchen_KG}} = 2.03, SD = .72$, alle $p < .001$). Demgegenüber zeigt sich auf der Komponente des benevolenten Sexismus kein bedeutsamer Unterschied zwischen Jungen der Experimental- und Kontrollgruppe ($M_{\text{Jungen_EG}} = 3.44, SD = .60, M_{\text{Jungen_KG}} = 3.32, SD = .53$), jedoch weisen Mädchen der Experimentalgruppe eine höhere Zustimmung zum benevolenten Sexismus im Vergleich zu Mädchen der Kontrollgruppe auf ($M_{\text{Mädchen_EG}} = 3.80, SD = .56, M_{\text{Mädchen_KG}} = 3.00, SD = 1.05$, post hoc: $t(40) = 3.11, p < .01, d = 0.94$). Die Abbildungen 8 und 9 fassen das Befundmuster noch einmal grafisch zusammen.

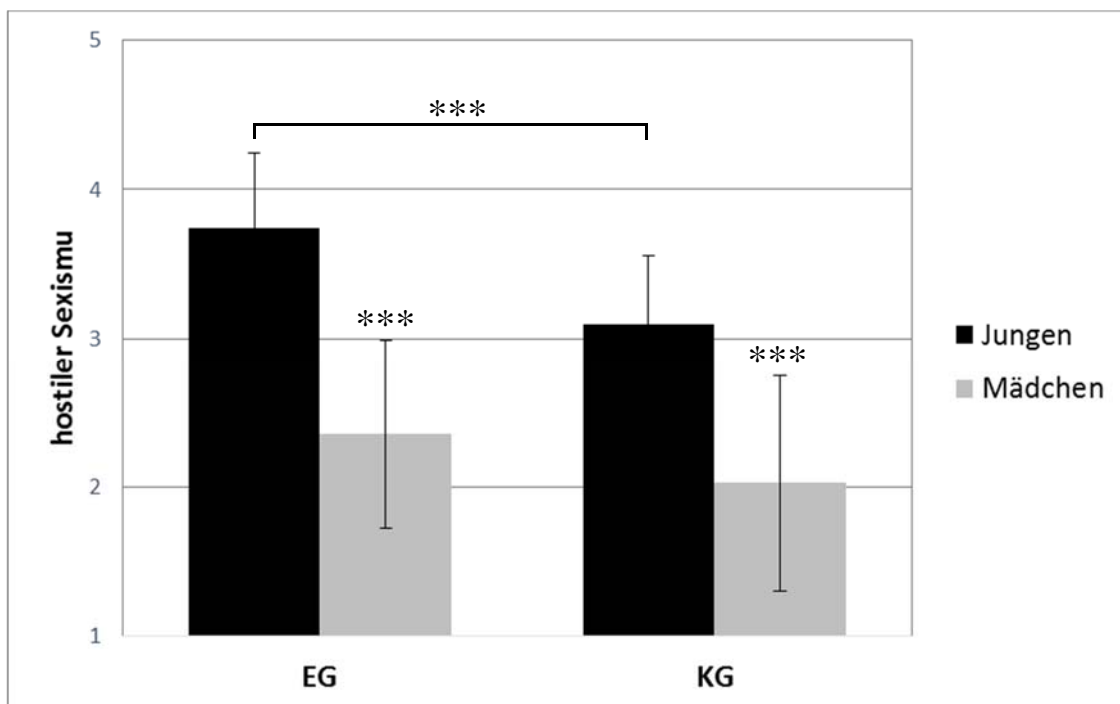


Abbildung 8: Zustimmung zu **hostilem** Sexismus nach Geschlecht und Untersuchungsbedingung.

* = $p < 0.05$, ** = $p < 0.01$ *** = $p < 0.001$

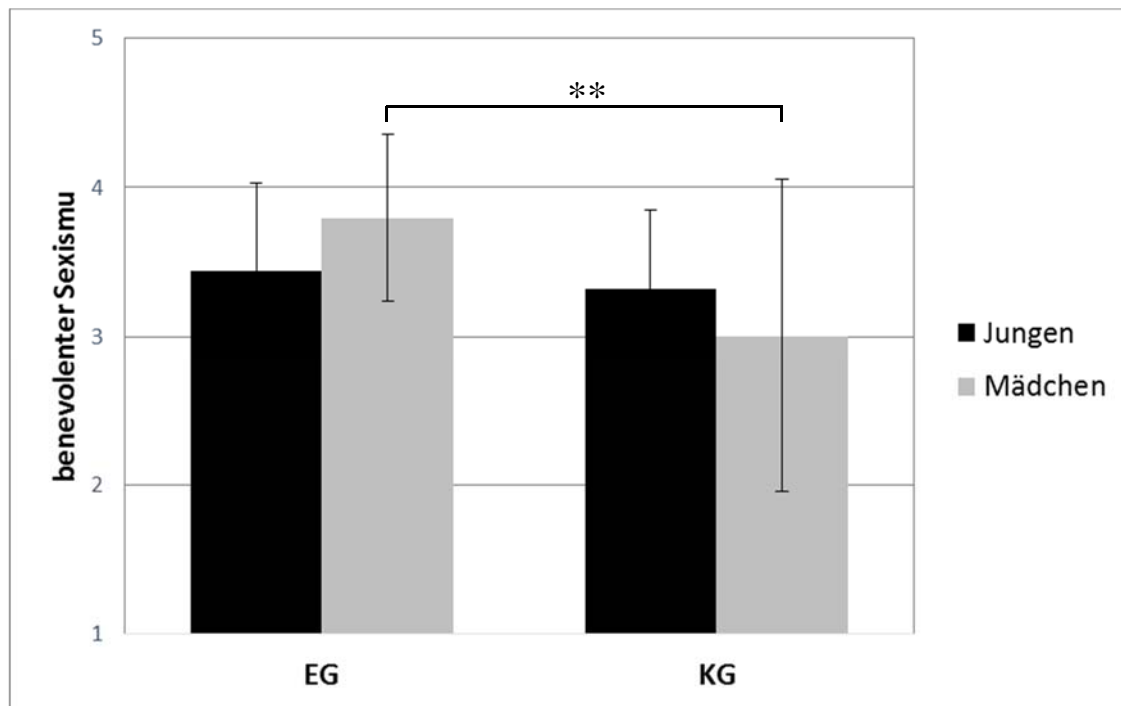


Abbildung 9: Zustimmung zu **benevolentem** Sexismus nach Geschlecht und Untersuchungsbedingung.

* = $p < 0.05$, ** = $p < 0.01$ *** = $p < 0.001$

Die Ergebnisse wurden wiederum kovarianzanalytisch überprüft, um mögliche Auswirkungen durch die Ungleichverteilung der ethnischen Herkunft und des sozioökonomischen Status zwischen den Geschlechtern bzw. Untersuchungsgruppen aus den Ergebnissen herauszupartialisieren. Die Ergebnisse zeigen, dass bis auf den zuvor signifikanten Haupteffekt auf dem Messwiederholungsfaktor (entspricht der hostile vs. benevolenten Komponente von ambivalentem Sexismus), welcher nunmehr mit $F < 1$ nicht mehr das geforderte Signifikanzniveau erreicht, die Ergebnisse der vorausgegangenen Varianzanalyse in Richtung und Stärke repliziert werden können (Interaktion Messwiederholungsfaktor \times Geschlecht: $F(1,70) = 40.73$, $p < .001$, $\eta^2 = .37$; Dreifachinteraktion Messwiederholungsfaktor \times Bedingung \times Geschlecht: $F(1,70) = 7.94$, $p < .01$, $\eta^2 = .10$).

In einer ergänzenden Studie wurde eine weitere, mögliche Strategie zur Aufwertung der Eigengruppe, resp. zur Abwertung der Außengruppe untersucht.

9. Studie 6: Ergänzungsstudie zur Abwertung von Weiblichkeit

Als eine weitere Strategie direkten sozialen Wettbewerbs wurde neben der in Studie 5 untersuchten rein hostile Abwertung von Weiblichkeit mittels ökonomischen Auszahlungsmatrizen (vgl. Simon & Trötschel, 2007) eine Möglichkeit sozialen Wettbewerbs betrachtet, bei welchem nicht nur die Außengruppe benachteiligt, sondern vielmehr auch die Eigengruppe bevorzugt werden kann.

Hinsichtlich der zur Bearbeitung vorgelegten Auszahlungsmatrizen muss zwischen symmetrischen und asymmetrischen Auszahlungsmatrizen unterschieden werden. Symmetrische Matrizen kennzeichnen sich dadurch, dass je mehr Ressourcen der eigenen Gruppe zugewiesen werden, desto weniger die andere Gruppe erhält, bei asymmetrischen Matrizen hingegen die *Outgroup* relativ mehr gewinnt, je mehr man der eigenen Gruppe zuweist. Im Rahmen dieser Ergänzungsstudie sollen die Annahmen geprüft werden, ob Jungen der Experimentalgruppe nach Bearbeitung von Materialien, welches Jungen als Bildungsverlierer darstellt, bei symmetrischen Auszahlungsmatrizen stärker in Richtung höherer Auszahlungen für die *Ingroup* vom Fairnesspunkt abweichen und bei asymmetrischen Auszahlungsmatrizen meist den Fairnesspunkt (die Mitte der Matrix) wählen, damit die *Outgroup* nicht relativ mehr gewinnt als die Eigengruppe. In Hypothesenform lauten die zu überprüfenden Annahmen für die Ergänzungsstudie somit wie folgt:

H13 Bei der Verteilung knapper Ressourcen mit einer symmetrischen Auszahlungsmatrix weisen Jungen in der Experimentalbedingung im Vergleich zu Jungen der Kontrollbedingung der *Ingroup* systematisch höhere Beträge zu.

H14 Bei der Verteilung knapper Ressourcen mit einer asymmetrischen Auszahlungsmatrix wählen Jungen in der Experimentalbedingung im Vergleich zu Jungen der Kontrollbedingung eher den Fairnesspunkt, damit die *Outgroup* nicht relativ mehr gewinnt.

Zusammenfassend soll untersucht werden, ob durch diese andere, direktere Form sozialen Wettbewerbs Jungen der Experimentalgruppe versuchen, den als illegitim wahrgenommenen Überlegenheitsstatus von Mädchen durch eine Bevorzugung im Sinne einer verstärkten Ressourcenzuweisung zugunsten der *Ingroup* zu begegnen.

9.1. Methode

Teilnehmende. Ergänzend zu den soeben erläuterten Ergebnissen wurde als eine weitere Strategie zur Aufwertung der *Ingroup* der Jungen an einer weiteren Stichprobe von $N = 103$ Schüler/-innen (45 Jungen und 58 Mädchen) der Klassenstufe 9 einer integrierten Sekundarschule untersucht. Auf die Experimentalbedingung entfielen 22 Jungen und 30 Mädchen, die Kontrollgruppe bestand aus 23 Jungen und 28 Mädchen. Das Alter der Stichprobe liegt im Mittel

bei $M = 15.10$ Jahren ($SD = 0.53$, $range = 14-17$ Jahre). Zwei von 45 Jungen und zwei von 52 Mädchen weisen eine direkte Migrationserfahrung auf (6 Mädchen ohne Angabe), unter Hinzunahme der mit den Eltern gesprochenen Sprache als Indikator für einen indirekten Migrationshintergrund geben vier von 41 Jungen und neun von 50 Mädchen an, mit ihren Eltern eine andere Sprache als Deutsch zu sprechen (4 Jungen und 8 Mädchen ohne Angaben), die Unterschiede sind mit $\chi^2 < 1.25$, $p > .26$ ohne statistische Bedeutung. Unter Berücksichtigung der Untersuchungsbedingung ergibt sich mit null von 49 im Ausland geborenen Personen in der Experimentalgruppe (3 Personen ohne Angaben) und 4 von 49 im Ausland geborenen Kontrollgruppenteilnehmer/-innen (2 Teilnehmer/-innen ohne Angabe) ein bedeutsamer Unterschied zwischen den Untersuchungsbedingungen, $\chi^2(1, N = 98) = 4.17$, $p = .04$, welcher unter Heranziehung der mit den Eltern gesprochenen Sprache mit fünf von 47 Personen in der Experimentalgruppe (5 fehlende Angaben) und 8 von 44 Personen in der Kontrollgruppe (7 fehlende Angaben) nicht aufrecht erhalten werden kann, $\chi^2(1, N = 91) = 1.06$, $p = .30$. Mit im Mittel $M = 5.92$ von 7 verfügbaren Gütern aus den sozioökonomischen Fragen ($SD = 1.38$, $range = 1-7$) liegt ein vergleichsweise hohes sozioökonomisches Niveau vor, welches zwischen Geschlecht bzw. Untersuchungsbedingung keine systematischen Unterschiede aufweist, $t(93) < |.60|$, $p > .55$.

Prozedur. Der Ablauf der Ergänzungsstudie war analog der Prozedur der Hauptuntersuchung. Auf eine mündliche Instruktion erfolgte die Präsentation und die Bearbeitung derselben Übungsaufgabe und anschließend der identischen Treatment-Aufgaben, d.h. die Interpretation einer Statistik über männliche Bildungsverlierer vs. den Wandel des Schulsystems in Hamburg und Bayern incl. Kontrollfragen zum Aufgabenverständnis. Abweichend zur Hauptstudie wurden allerdings keine Vignetten vorgelegt, sondern Auszahlungsmatrizen (vgl. Abb. 10 und 11) zur Verteilung knapper Ressourcen eingebettet in eine instruierende Geschichte.

Gruppe der Jungen →		10 €	20 €	30 €	40 €	50 €	60 €	70 €	80 €	90 €	100 €	110 €	120 €	130 €	140 €	150 €
	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Gruppe der Mädchen →	150 €	140 €	130 €	120 €	110 €	100 €	90 €	80 €	70 €	60 €	50 €	40 €	30 €	20 €	10 €	

Abbildung 10: symmetrische Auszahlungsmatrix

		Fairnesspunkt											
Gruppe der Jungen →	70 €	80 €	90 €	100 €	110 €	120 €	130 €	140 €	150 €	160 €	170 €	180 €	190 €
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gruppe der Mädchen →	10 €	30 €	50 €	70 €	90 €	110 €	130 €	150 €	170 €	190 €	210 €	230 €	250 €

Abbildung 11: asymmetrische Auszahlungsmatrix¹⁰

Die Schüler/-innen wurden dabei gebeten, fiktive Geldbeträge zuzuweisen, wobei zuerst eine symmetrische, 16-stufige Auszahlungsmatrix bearbeitet wurde, bei der mit jedem Mehrbetrag, welcher der *Ingroup* zugewiesen wird, die *Outgroup* relativ gesehen mehr Geldbeträge verliert. Im Fortgang der Instruktionsgeschichte war zusätzlich ein zweite, diesmal aber asymmetrische Matrix mit 13 Stufen zu bearbeiten. Bei dieser befindet sich in der Mitte der Skala (Stufe 7) ein sog. Fairnesspunkt (Zuweisung gleicher Geldbeträge an *In-* und *Outgroup*), bei Überschreitung dieses Punktes gewinnt die Außengruppe relativ zur Eigengruppe größere Geldbeträge. Durch ankreuzen einer bestimmten Betragskonstellation auf der Matrix wurden die zugewiesenen Geldbeträge abgefragt. Den Abschluss der Untersuchung mit einer Dauer von ca. 45 Minuten bildete die Abfrage der soziodemografischen Daten und eine sorgfältige Aufklärung aller Teilnehmenden.

9.2. Ergebnisse

Die Analyse der Anzahl korrekter Antworten bezüglich der Treatmentfragen ergibt, dass im Mittel $M = 56.67\%$ ($SD = 13.19\%$, *range* 50-100%) der Fragen zur Manipulationsüberprüfung richtig beantwortet wurden, was einem ausreichenden Anteil korrekt beantworteter Kontrollfragen entspricht. Zwischen Jungen und Mädchen ergeben sich mit $M_{\text{Jungen}} = 57.50\%$ ($SD = 13.95\%$) und $M_{\text{Mädchen}} = 56.03\%$ ($SD = 12.66\%$) keine bedeutsamen Unterschiede hinsichtlich der Anteile korrekter Antworten, $t(101) = .56$, $p = .58$. Jedoch liegt zwischen den Untersuchungsbedingungen mit $M_{\text{EG}} = 60.82\%$ ($SD = 16.06\%$) und $M_{\text{KG}} = 52.45\%$ ($SD = 7.51\%$) ein signifikanter Unterschied in Richtung eines höheren Anteils korrekter Antworten in der Experimentalbedingung vor, $t(101) = 3.40$, $p < .001$, $d = 0.67$. Wie auch in der Hauptstudie kann dieser bedeutsame Unterschied im Kontext der vorliegenden Studie unterstützend wirken, da für das Gelingen der Untersuchung die korrekte Verarbeitung der Treatmentmaterialien für die Experimentalgruppe von höherer Bedeutung ist, zumal auch in der Kontrollgruppe ein ausreichendes Niveau an korrekt beantworteten Kontrollfragen festgestellt werden kann.

Die auf der symmetrischen vs. asymmetrischen Auszahlungsmatrix zugewiesenen Beträge wurden einer 2 (Bedingung) \times 2 (Geschlecht) \times (2) (Messwiederholungsfaktor: symmetrische vs. asymmetrische Auszahlungsmatrix) *mixed model*-ANOVA unterzogen. Die Ergebnisse

¹⁰ Die Markierung des Fairnesspunkts dient lediglich der Illustration und war in der Version, welche den Schüler/-innen zur Bearbeitung vorgelegt wurde **nicht** enthalten.

indizieren einen signifikanten Haupteffekt auf dem Messwiederholungsfaktor in Richtung einer Zuweisung der Beträge eher zugunsten von Jungen auf der symmetrischen Auszahlungsmatrix (höhere Werte bezeichnen eine höhere Ressourcenzuweisung an Jungen), $M = 8.47$ ($SD = 2.89$), wohingegen auf der asymmetrischen Auszahlungsmatrix eher der Fairnesspunkt gewählt wird, $M = 7.31$ ($SD = 2.88$), $F(1,93) = 9.83$, $p < .01$, $\eta^2 = .10$. Ein genaueres Bild von der Befundlage liefert die signifikante Interaktion aus dem Messwiederholungsfaktor und dem Faktor Geschlecht. Während Jungen generell, ohne Einfluss der Untersuchungsbedingung, mit $M_{\text{Jungen}} = 9.23$ ($SD = 3.35$) auf der symmetrischen Auszahlungsmatrix höhere Beträge an die Eigengruppe vergeben, als Mädchen mit $M_{\text{Mädchen}} = 7.85$ ($SD = 2.30$), liegen die zugewiesenen Ressourcen beider Geschlechter auf der asymmetrischen Auszahlungsmatrix nahe des Fairnesspunkts, $M_{\text{Jungen}} = 6.77$ ($SD = 2.87$), $M_{\text{Mädchen}} = 7.75$ ($SD = 2.85$), Interaktionsterm: $F(1,93) = 8.09$, $p < .01$, $\eta^2 = .08$. Post hoc-Einzelvergleiche unterstützen diesen Befund, indem mit $t(95) = 2.31$, $p = .01$, $d = 0.48$ lediglich die Unterschiede zwischen Jungen und Mädchen auf der symmetrischen Matrix statistische Bedeutsamkeit erlangen, nicht jedoch auf der asymmetrischen Matrix, $t(96) = -1.70$, $p = .09$. Sowohl die Interaktion aus dem Messwiederholungsfaktor und dem Faktor Untersuchungsbedingung ($F(1,93) = 2.57$, $p = .11$) als auch die Dreifachinteraktion ($F < 1$, n.s.) liegen oberhalb der zur Interpretation notwendigen Irrtumswahrscheinlichkeit.

9.3. Gesamtdiskussion Studien 5 & 6

Die vorliegende Untersuchung diene der Feststellung, inwiefern eine Vorhersage zum Bedürfnis der (Wieder-) Herstellung positiver Distinktheit zwischen Eigen- und Fremdgruppe nach der *Social Identity Theory* (Tajfel & Turner, 1986) auf das Problem „Jungen als Bildungsverlierer“ übertragen werden kann. Genauer wurde dies in verschiedenen experimentellen (Teil-) Studien untersucht, welche sich als Wiederherstellung positiver Distinktheit durch Eigengruppenaufwertung und Fremdgruppenabwertung subsumieren lassen. Es findet sich insgesamt eine gemischte Evidenz, welche in die Richtung *keiner* systematischen Aufwertung der Eigengruppe, jedoch einer hostilen Abwertung der Fremdgruppe deutet.

In einem ersten Schritt wurde überprüft, ob im Gegensatz zum stark feminin konnotierten Kontext Schule (Hannover & Kessels, 2011) die Inszenierung von Maskulinität eine adäquate Möglichkeit bereitstellt, sich einer bedrohten sozialen Identität rückzuversichern. Die hierzu durchgeführte Vignettenuntersuchung zeigt, dass die Ergebnisse nicht systematisch in Abhängigkeit der Experimentalbedingung zwischen den Geschlechtern variieren. Unabhängig von Geschlecht und Experimentalbedingung wurde die stark männliche Vignette als wärmer/emotional näher bewertet als die zweite, wenig männliche Vignette. Zwischen den Geschlechtern lässt sich bei Jungen auf der Skala emotionale Nähe eine leichte Eigengruppenfavorisierung im Sinne einer tendenziell besseren Bewertung beider Vignetten durch Jungen feststellen. Auf der zweiten ge-

messenen Skala, der wahrgenommenen Kompetenz der auf den Vignetten dargestellten Jungen, konnten keinerlei durch das Experiment induzierte Unterschiede festgehalten werden. Insbesondere durch den mangelnden Einfluss des experimentellen Treatments sind die Hypothesen 9 und 10 zum differenziellen Einfluss der Untersuchungsbedingung auf Jungen somit abzulehnen. In einem weiteren Schritt der Untersuchung wurde als zusätzliche Strategie die Abwertung der Fremdgruppe der Mädchen durch feindselige vs. benevolent-sexistische Verhaltensweisen untersucht. Die Ergebnisse aus der Befragung zu sexistischen Einstellungen machen erwartungskonform deutlich, dass Jungen unabhängig von der Untersuchungsbedingung der Komponente des hostilen (feindseligen) Sexismus generell stärker zustimmen als Mädchen. Weiterhin kann mit substantiell starken Effektgrößen festgestellt werden, dass Jungen der Experimentalgruppe im Vergleich zu Jungen der Kontrollgruppe deutlich stärker hostile ausgeprägte Facetten von Sexismus zeigen, während sich Mädchen zwischen den Untersuchungsbedingungen nicht in ihren hostile-sexistischen Einstellungen unterscheiden. Die Hypothesen 11 und 12 sind somit beizubehalten. Zusätzlich wurde festgestellt, dass Mädchen der Experimentalbedingung der benevolenten Komponente von Sexismus stärker zustimmen als Mädchen der Kontrollgruppe. Dieses Ergebnismuster, welches so nicht erwartet wurde, kann dahingehend interpretiert werden, dass im Jugendalter in einem vergleichsweise frühen Entwicklungsstadium von Sexismus und in welchem nur sehr wenige Erkenntnisse zur Ausprägung und Entwicklung beider Komponenten vorliegen, der eigentlich sexistische und damit abwertende Anteil von Benevolenz durch Mädchen u.U. (noch) nicht erkannt und ausreichend reflektiert wird. Dies bewirkt eine verstärkte Zustimmung zu benevolent-sexistischen Fragestellungen, welche stark auf den Vorteil von Frauen aus benevolentem Sexismus fokussiert und wie eine „Belohnung“ für das vergleichsweise bessere Abschneiden von Mädchen in der Schule interpretiert werden kann. Alternativ kann argumentiert werden, dass auch benevolenter Sexismus die Funktion der Aufrechterhaltung von Statusunterschieden zwischen den Geschlechtern erfüllt (Fiske & Glick, 1995). Damit erscheint es möglich, dass Mädchen benevolenten Sexismus bewusst nutzen, um ambivalent sexistische Einstellungen gegenüber Jungen auszudrücken. Innerhalb des Systems, in dem Mädchen weniger Macht und Status haben, können Mädchen die Fremdgruppe der Jungen in solchen Bereichen manipulieren, in denen Jungen von ihnen abhängig sind, und auf diese Weise Macht ausüben und darüber einen gewissen Status erreichen (Rau, 2013). Dieser Interpretation folgend würde die Aktivierung des Stereotyps vom männlichen Schulversager nicht nur hostilem Sexismus von Jungen gegenüber Mädchen, sondern auch von Mädchen gegenüber Jungen verstärken. Bezüglich der Komponente des benevolenten Sexismus, zu welcher keine Hypothesen formuliert wurden, ist somit ein differenzieller Einfluss des experimentellen Treatments auf Mädchen festzuhalten, dessen genaue Ursache aus den vorliegenden Daten nicht festzustellen ist und deren Aufdeckung einer weiteren Untersuchung vorbehalten bleibt.

In einer weiteren Studie wurden beide Möglichkeiten – Aufwertung der *Ingroup* und Abwertung der *Outgroup* – in eine gemeinsame Aufgabe integriert, um beide möglichen Strategien innerhalb einer Aufgabenstellung gemeinsam untersuchen zu können. Bei den hierfür eingesetzten Auszahlungsmatrizen zeigt sich, dass Jungen auf der symmetrischen Auszahlungsmatrix unabhängig von der Experimentalbedingung eine Eigengruppenfavorisierung im Sinne einer stärkeren Ressourcenzuweisung für die *Ingroup* zeigen. Auf der asymmetrischen Auszahlungsmatrix wird von beiden Geschlechtern unabhängig vom experimentellen Treatment eher der Fairnesspunkt (ca. gleich hohe Auszahlungen an die Eigen- und Fremdgruppe) gewählt. Mangels eines Einflusses der Experimentalbedingung sind somit Hypothesen 13 und 14 abzulehnen. Insgesamt liefert die Ergänzungsstudie somit zusätzliche Evidenz für eine Eigengruppenfavorisierung der Gruppe von Jungen, jedoch unabhängig vom zuvor präsentierten Treatmentmaterial: Möglicherweise ist die untersuchte Altersgruppe (Jugendalter) generell motiviert, eine möglichst positive Distinktheit der eigenen Geschlechtergruppe zu erreichen, ohne dass durch das applizierte Treatment ein verstärkender Effekt induziert werden konnte.

In einer Untersuchung zur Anwendung der Theorie der sozialen Identität wurden bei Ahmed (2007) ökonomische, dyadische Spiele aus der Spieltheorie (z.B. das Gefangenendilemma) im Rahmen eines *minimal group*-Paradigmas untersucht. Der Autor findet eine deutliche Diskriminierung der Außengruppe, in der aus je einem Mitglied der Eigen- und Fremdgruppe bestehenden Dyaden, im Vergleich zu *Ingroup-Ingroup*-Dyaden oder einer Kontrollgruppe. Zusätzlich fand keinerlei hostile Abwertung der Außengruppe statt, sondern ausschließlich eine Diskriminierung durch Favorisierung der eigenen Gruppe, was bereits Brewer (1979) in einem Review belegte und auch in Bezug auf hostile, aggressive Reaktion bestätigt wurde (Struch & Schwartz, 1989). Im Kontext der vorliegenden Untersuchung können diese Befunde neben möglichen methodischen Einschränkungen (vgl. Limitationen) unter Hinzunahme dieses empirischen Befundes zur Erklärung des generellen *ingroup bias* bei Jungen beitragen, stehen allerdings im Gegensatz zu den hostile Abwertungen von Jungen ggü. Mädchen hinsichtlich deren sexistischen Einstellungen. Da auch Jungen der Kontrollgruppe eine stärker hostile sexistische Einstellung aufwiesen als Mädchen, bleibt an dieser Stelle allerdings zu diskutieren, inwiefern hostile sexistische Einstellungen nicht eine Strategie darstellen, welche von Jungen generell angewendet wird, um eine positive Distinktheit der Eigengruppe zu erreichen oder abzusichern. Es scheint so, als würden Jungen der Experimentalgruppe getriggert durch die Bedrohung der eigenen Identität durch die verstärkte Ausprägung eines hostile Sexismus eine positive Distinktheit wiederherzustellen versuchen, wohingegen Jungen der Kontrollgruppe die gleiche Strategie in abgemilderter Form zur Stabilisierung einsetzen. Es bleibt zu erklären, warum diese Effekte auf den anderen Maßen (Vignetten / Auszahlungsmatrizen) nicht in Richtung dieser Interpretation festzustellen sind. Roccas & Schwartz (1993) weisen auf die zentrale Rolle der Valenz der

Dimensionen für den Gruppenvergleich hin. Es konnte gezeigt werden, dass Eigengruppenfavorisierungen nur auf als relevant wahrgenommenen Dimensionen oder Verhaltensweisen auftreten. Im Kontext der vorliegenden Studie kann dieser Befund u.U. ergänzend erklären, warum auf dem direktesten Verhaltensmaß, dem Sexismus, differenzielle Ergebnisse festgestellt werden konnten, wohingegen diese auf weniger direkten Maßen (*fiktive* Vignettenbewertungen und *fiktive* Ressourcenzuweisungen) ausblieben. Es könnte somit unter Rückgriff auf Roccas & Schwartz (1993) argumentiert werden, dass Sexismus ein relevanteres und verhaltensnäheres Konstrukt darstellt, da es näher am Kernkonzept Maskulinität liegt als die anderen erhobenen abhängigen Variablen und damit am ehesten selektiv als Reaktion auf eine Bedrohung der geltenden Intergruppenverhältnisse gezeigt wird.

Die (mit Ausnahme der Ergebnisse bezüglich des Sexismus) fehlenden Effekte des Treatments werden weiterhin durch Befunde an einer jüngeren Altersgruppe von Kindern unterstützt: Einerseits wird durch experimentell induzierte, minimale Gruppen eine Eigengruppenfavorisierung ausgelöst, wobei jedoch mit vergleichsweise stärkeren Effekten eine Verzerrung zugunsten der eigenen Geschlechtergruppe (*gender bias*) dominiert (Dunham, Baron & Carey, 2011). Konträr stellt Spielman (2000) gegenüber, dass ohne kompetitives *Priming* bei Kindern keine Eigengruppenfavorisierung festgestellt wurde. Die vorliegende Untersuchung ergänzt um eine weitere Altersspanne, das Jugendalter, und unterstützt ausschnittsweise beide Befunde. Hinsichtlich des ambivalenten Sexismus hat ein kompetitives *Priming* die Eigengruppenfavorisierung (Jungen durch hostile Abwertung der Fremdgruppe, Mädchen durch eine wahrgenommene Aufwertung der Eigengruppe *via* benevolentem Sexismus) deutlich verstärkt, wohingegen bei der ergänzenden Studie (Auszahlungsmatrizen) ein kompetitiver *Prime* keine notwendige Bedingung für eine Eigengruppenfavorisierung darstellte. Damit weisen die Ergebnisse der vorliegenden Studie auf die Notwendigkeit hin, nicht nur nach experimentell induzierten Gruppen zu differenzieren, sondern auch natürlich bestehende Gruppen stärker mit zu berücksichtigen, insbesondere dann, wenn Geschlecht wie im Jugendalter eine besonders saliente Dimension darstellt. In einer Studie von Jetten, Spears & Manstead (1996) konnte diesbezüglich gezeigt werden, dass der Einfluss von Intergruppennormen bei der Zuweisung von knappen Ressourcen zwischen experimentellen und natürlichen Gruppen divergiert.

Betrachtet man die verschiedenen Ergebnisse, insbesondere im Hinblick auf die eingesetzte Strategie (*Aufwertung* der Eigengruppe vs. *Abwertung* der Fremdgruppe), wird deutlich, dass die Aufwertung der Eigengruppe bei Jungen unabhängig vom experimentellen Treatment erfolgt, wohingegen die Abwertung der Fremdgruppe einen durch das experimentelle Treatment verursachten, differenziellen Effekt aufweist. Zu den motivationalen Grundlagen beider Strategien können aus der *Social Identity Theory* (Tajfel & Turner, 1986) zwei Prämissen abgeleitet werden: Es wird postuliert, dass a) ein bedrohter Selbstwert eine Diskriminierung der

Fremdgruppe motiviert sowie b) eine erfolgreiche Diskriminierung der Fremdgruppe – egal auf welche Art und Weise – den Selbstwert wieder erhöht. Rubin & Hewstone (1998) finden allerdings in einem ausführlichen Review zahlreicher Untersuchungen kaum Evidenz für ein Bedürfnis nach Diskriminierung, welches durch eine Selbstwerterhöhung motiviert wird, wenngleich zahlreiche Belege für eine Erhöhung des Selbstwertes durch diskriminierendes Intergruppenverhalten existieren. Diese Erkenntnisse sind im Bezug zur aktuellen Studie nur partiell in Bezug zu setzen, da keine selbstwertbezogenen Daten mit erfasst wurden. Die differenziellen Ergebnisse, insbesondere zum hostilem Sexismus, deuten jedoch darauf hin, dass die stärkere Feindseligkeit bei Jungen in der Experimentalgruppe psychologisch durch eine Motivation zur Selbstwerterhöhung angestoßen wird. Gegen eine selbstwerterhöhende Motivation sprechen sowohl die Ergebnisse der aufbauenden Studie als auch die Tatsache, dass auch Jungen der Kontrollgruppe stärker hostile sexistische Einstellungen berichten als Mädchen (und ferner Mädchen – deren Selbstwert zu keinem Zeitpunkt der Untersuchung bedroht war – in der Experimentalbedingung stärker benevolent sexistische Einstellungen aufweisen). Zusammenfassend scheint es so zu sein, dass eine selbstwerterhöhende Motivation diskriminierendes Intergruppenverhaltens zumindest nicht generell abgelehnt werden kann, vielmehr ist diese differenzierter nach Eigengruppenfavorisierung oder hostiler Fremdgruppenabwertung sowie Art und Stärke der Bedrohung einer sozialen Gruppe zu betrachten. Die hier unter Einbezug der genannten empirischen Arbeiten getroffenen Aussagen stellen allerdings nur erste Vermutungen dar und müssten durch weitere Forschung einer genaueren Prüfung zugeführt werden – eine Ableitung konkreter Schlussfolgerungen ist aus den vorliegenden Daten nicht möglich.

Limitationen. Es ist denkbar, dass die beschriebenen, divergierenden Befunde hinsichtlich der Aufwertung der Eigengruppe (Auszahlungsmatrizen) vs. Abwertung der Fremdgruppe (Sexismus) durch die Operationalisierung der abhängigen Variablen – insbesondere im Hinblick auf die Auszahlungsmatrizen – entstanden sind. Der Umgang mit den Auszahlungsmatrizen erfordert ein hohes Maß an ökonomischer Entscheidungsfähigkeit, dies betrifft im Besonderen die asymmetrische Matrix. Es kann nicht ausgeschlossen werden, dass bei der Bearbeitung der Matrizen trotz einer sorgfältigen Instruktion und einer einbettenden Erzählung Verständnisprobleme aufgetreten sind, die sich auf die daraus resultierende Entscheidung für die Verteilung knapper Ressourcen ausgewirkt haben. Während bei der symmetrischen Matrix deutlich gemacht werden konnte, dass ein und derselbe Betrag zwischen der Gruppe von Jungen und Mädchen aufgeteilt werden sollte (d.h. der Gesamtbetrag der zu verteilenden Ressourcen blieb in der Summe stets unverändert) variierten die Gesamtbeträge auf der asymmetrischen Matrix deutlich, was Verwirrungen begünstigt haben kann. Es kann damit nicht abschließend geklärt werden, ob die divergierenden Befunde durch Schwierigkeiten bei der Bearbeitung der Instrumente zur Erfassung der abhängigen Variablen begünstigt wurden. In zukünftigen Untersuchungen

wäre daher anzuraten, im Rahmen eines Vortests festzustellen, ob hinreichende Kompetenzen für die Bearbeitung spezieller Erfassungsinstrumente jenseits typischer Fragebogenmaße im Alters- und Kenntnisbereich der angedachten Zielstichprobe vorhanden sind (für ökonomische Entscheidungen vgl. Gummerum, Keller, Takezawa & Mata, 2008; Schug & Birkey, 1985). Darüber hinaus wäre denkbar, die Plastizität und Verständlichkeit dadurch zu erhöhen, dass für ökonomische Entscheidungen vermehrt Anschauungsmaterial (wie z.B. „Spielgeld“ zur Imitation einer realen Geldvergabe an zwei Gruppen) eingesetzt wird, wobei zusätzlich typische (ökonomisch relevante) Heuristiken und Urteilsverzerrungen (für einen Überblick: Raab, Unger & Unger, 2010, Kap. 9; Strack & Deutsch, 2002) berücksichtigt werden sollten.

Ferner betrachtet die vorliegende Studie ausschließlich einen direkten Einfluss eines identitätsbedrohenden Reizes auf eigengruppenbezogene Aufwertungsprozesse sowie (hostile) Abwertungen der Fremdgruppe. Verschiedene Arbeiten, welche die Theorie der sozialen Identität mit einem Schwerpunkt auf den *ingroup bias* angewendet haben, weisen allerdings darauf hin, dass verschiedene, vermittelnde Mechanismen beim Zustandekommen einer Eigengruppenfavorisierung wirksam sind. Jackson & Smith (1999) stellen fest, dass die soziale Identität ein multidimensionales Konstrukt ist und in zwei Basistypen aufgeteilt werden kann: Einerseits in eine unsichere Identität, welcher andererseits eine gesicherte soziale Identität gegenüber steht. Ausschließlich dann, wenn eine ungesicherte Identität vorliegt, kann den Ergebnissen folgend eine Favorisierung der Eigengruppe und ferner eine erhöhte Wahrnehmung von Eigengruppenhomogenität erwartet werden. Durch den experimentellen Ansatz der vorliegenden Studie kann erwartet werden, dass ungesicherte soziale Identitäten induziert werden, was die Herausbildung eines differenziellen *ingroup bias* wahrscheinlich macht, allerdings im Gegensatz zu den festgestellten Befunden steht. Eine genauere Überprüfung, ob tatsächlich eine ungesicherte soziale Identität durch das Treatment vermittelt wurde, könnte zu einer weiteren Klärung der o.g. Befunde beitragen. Ferner spielt die wahrgenommene Relevanz der Vergleichsdimensionen und damit einhergehend auch die Relevanz der im Treatment dargebotenen Informationen eine wichtige Rolle (Roccas & Schwartz, 1993), das heißt, es wäre u.U. von Vorteil, genauer zu erfassen, welche Relevanz dem Kontext Schule und schulische Leistung durch Jungen zugeschrieben wird, um die Wirksamkeit des Treatments als Mittel zur Induktion einer ungesicherten sozialen Identität noch stärker abzusichern.

Zusätzlich wäre es erstrebenswert, durch eine ausgewogenere Auswahl der Fragen zur Manipulationsüberprüfung die gefundenen Leistungsunterschiede bei den Kontrollfragen zu vermeiden. Zwar wurde das Material zur Bedrohung der sozialen Identität insbesondere durch die Zielgruppe der Jungen besser verarbeitet, jedoch sollte auch bei den Vergleichsgruppen eine ebenso sorgfältige Verarbeitung der jeweiligen Treatment- oder Kontrollmaterialien gewährleis-

tet werden, um eine solide *Baseline* sicherzustellen (auch wenn eine Verarbeitung der jeweiligen Materialien auf ausreichendem bis zufriedenstellendem Niveau abgesichert wurde).

Mlicky & Ellemers (1996) weisen zusätzlich darauf hin, dass eine Distinktheit von der Fremdgruppe nicht notwendigerweise von positiver Valenz sein muss, sondern auch eine negative Distinktheit mit dem Ziel erreicht werden kann, sich zumindest von einer anderen Gruppe deutlich abzugrenzen. Für den Kontext „Jungen als Bildungsverlierer“ und die vorliegende Untersuchung bedeutet dies, dass auch negativ konnotierte Verhaltensweisen (wie in den präsentierten Vignetten) oder Verhaltensweisen (wie Sexismus) geeignet sein können, eine Stabilisierung der sozialen Identität zu erreichen. Es kann diskutiert werden, ob den Jugendlichen in der vorliegenden Untersuchung die gesellschaftliche Wahrnehmung und die daraus resultierenden Konsequenzen von Sexismus (noch) nicht bewusst sind oder aber diese im Sinne von Mlicky & Ellemers (1996) billigend und wider der negativen, gesellschaftlichen Wahrnehmung eingesetzt werden, um sich von der Außengruppe der Mädchen/Frauen abzugrenzen. Der Befund spricht zumindest dafür, dass das in der vorliegenden Studie eingesetzte Material eine Möglichkeit zur Absicherung sozialer Identität anbietet und impliziert, dass verschiedene Valenzen von Distinktheit im Rahmen weiterer Untersuchungen entsprechend berücksichtigt werden sollten.

Weiterhin sei darauf hingewiesen, dass insbesondere zur Abklärung des Befundes zu höheren benevolent-sexistischen Einstellungen von Mädchen nach der Darstellung von Jungen als Verlierer in der Schule im Vergleich zur Kontrollgruppe eine Ergänzung der Untersuchung um Skalen zur Erfassung von sexistischen Einstellungen gegenüber Jungen von zwingender Notwendigkeit wäre. Mit den aktuell verfügbaren Erfassungsmethoden (z.B. Werner & Collani, 2002) scheint diese Forderung für das Jugendalter zum gegebenen Zeitpunkt nicht realisierbar, da diese Skalen eine substantielle Beziehungserfahrung einhergehend mit der Erkenntnis gegenseitiger Abhängigkeiten in einer Beziehung voraussetzen, was in der erfassten Altersspanne nicht erwartet werden kann. Wäre eine Erhebung von ambivalentem Sexismus gegenüber Jungen möglich, so wären bei Mädchen möglicherweise höhere hostile Sexismus-Werte in Reaktion auf die Darstellung von schulversagenden Jungen, insbesondere Dimensionen wie Macht und Maternalismus, zu erwarten. Abschließend erscheint es zielführend und im Hinblick auf die aktuelle Befundlage geboten, in zukünftigen Untersuchungen zu Auf- und Abwertungsprozessen im Intergruppenverhalten den kollektiven Selbstwert (Branscombe & Wann, 1994; Crocker & Luhtanan, 1990), den Wert der Eigengruppe und den individuellen Selbstwert (Platow, Harley, Hunter, Hanning, Shave & O'Connell, 1997), die Identifizierung mit der Eigengruppe (Branscombe & Wann, 1994; Gagnon & Bourhis, 1996) oder die wahrgenommene Prototypikalität des Individuums für die Eigengruppe (Jetten, Spears & Manstead, 1997) mit zu erfassen. In der Gesamtheit weisen alle genannten Arbeiten darauf hin, dass eine höhere wahrgenommene Nähe des Individuums zur Eigengruppe (im Sinne einer höheren *group identification*, eines höheren sub-

jektiven Wertes der eigenen Gruppe, einer Wertung der eigenen Person als prototypisch für die Eigengruppe, aber auch eines wahrgenommenen hohen kollektiven Selbstwertes) die Vulnerabilität gegenüber einer Bedrohung der sozialen Identität maßgeblich erhöht bzw. bei geringeren Ausprägungen auf den genannten Einflussgrößen zumeist keinerlei Auswirkung einer sozialen Identitätsbedrohung nachgewiesen werden konnten. Dies impliziert, dass diese potenziell moderierenden oder mediierenden Variablen in Folgeuntersuchungen stärker einzubeziehen sind.

10. Studie 7: Wiederherstellung positiver Distinktheit II

Wechsel identitätsrelevanter, sozialer Vergleichsdimensionen als Bewältigungsstrategie auf das Stereotyp des männlichen Schulversagers.

10.1. Überblick und Fragestellung

In der im Folgenden vorgestellten Studie soll eine weitere Strategie zur Herstellung positiver Distinktheit für die Gruppe der Jungen im Sinne der *Social Identity Theory* (Tajfel & Turner, 1986) untersucht werden. Von Interesse ist hier, ob Jungen nach der Exposition des Stereotyps des männlichen Schulversagers entsprechend der Theorie die Vergleichsdimension für die Intergruppenvergleiche wechseln und somit den sozialen Vergleich (für einen Überblick: Mussweiler, 2003) nach Darstellung ihres geringeren Schulerfolgs auf anderen, nicht-schulischen Dimensionen anstellen, sodass die Eigengruppe wieder einen Überlegenheitsstatus einnimmt.

Für die Heranziehung einer alternativen Vergleichsdimension ist es erforderlich, dass die Eigengruppe auf dieser als erfolgreicher wahrgenommen wird als die Fremdgruppe. Zwar übersieht das medial transportierte Stereotyp, welches von einem generellen Versagen von Jungen in der Schule ausgeht, dass Jungen auch im schulischen Bereich keinesfalls die generellen Verlierer sind (sichtbar z.B. in besseren Kompetenzen in der Mathematik und im Problemlösen; OECD, 2014), jedoch ist den Ergebnissen der *Studie 1* folgend davon auszugehen, dass diese Unterschiede zugunsten von Jungen im schulischen Kontext nicht umfassend antizipiert werden. Weiterhin kann festgestellt werden, dass Mädchen zwar in der Schule bessere Abschlusszertifikate erwerben, als vergleichsweise Jungen (Statistisches Bundesamt, 2014a), und Mädchen auch als geeigneter für Schule wahrgenommen werden (*Studie 1*; auch: Heyder & Kessels, 2013), jedoch keine Umsetzung der besseren Abschlusszertifikate in entsprechendem Berufserfolg im Sinne von Vollbeschäftigung, Entlohnung und beruflicher Positionierung erfolgt (Holst & Wiemer, 2010; Leuze & Strauß, 2009). Entsprechend des Inhalts der gängigen Geschlechterstereotypen existieren Hinweise darauf, dass Männer typischerweise als *agentic* und Frauen als *communion* beschrieben werden, was zur Erklärung dieses Wandels von Geschlechterdifferenzen von der Schule ins Erwerbsleben beitragen kann (Abele, 2003; Diekman & Eagly, 2000; Madera, Hebl & Martin, 2009). Hieraus resultiert, dass maskulines Verhalten typischerweise mit Leitungsfunktionen und einer Hervorhebung analytischer Fähigkeiten assoziiert wird, wohingegen typisch feminines Verhalten eher durch empathisch-soziale Fähigkeiten gekennzeichnet ist. Dieser Wandel von der schulischen Überlegenheit der Mädchen hin zu einer Benachteiligung in der Erwerbsarbeit kann sozialpsychologisch damit erklärt werden, dass die Geschlechterdifferenzen in schulischer Leistung und schulischem Engagement nicht einen höheren Erfolg im Berufsleben determinieren, sondern vielmehr die Wahrnehmung einer besseren Passung außerschulischer und sozialer Kompetenzen (wie Führungsverhalten) von Jungen in der Gesellschaft, deren Berufserfolg positiv beeinflusst (Rudman & Glick, 1999) und ferner begünstigt dass die gesell-

schaftliche Wahrnehmung von Frauen als *communion* dazu führt, dass Frauen mehr mit Heim- und Erziehungsarbeit belastet werden (Heisig, 2011). Als eine weitere Facette als Quelle von Maskulinität wird die Betonung von körperlicher Stärke und Sportlichkeit angesehen (Bhana, 2008; Eccles & Harold, 1991; Ricciardelli, McCabe & Ridge, 2006; Schmalz & Kerstetter, 2006), was wiederum mit einer Zuschreibung weiterer, positiver Fähigkeiten wie Unabhängigkeit und Zuschreibung von Führungseigenschaften einhergeht (Larson, 2000). Den mit beruflichem Erfolg und körperlicher Leistungsfähigkeit/Sportlichkeit assoziierten Verhaltensweisen, welche eng an agentische Eigenschaften jenseits schulischer Leistungsfähigkeit wie u.a. Führungsqualität und Durchsetzungsfähigkeit gekoppelt sind, scheint also eine besondere Bedeutung bei der Frage zuzukommen, auf welchen Vergleichsdimensionen Jungen jenseits schulischer Leistung wahrnehmbar besser abschneiden als Mädchen. Es kann somit vermutet werden, dass Jungen nach der Bedrohung der Eigengruppe auf den Dimensionen Leistungsfähigkeit und Engagement im Kontext Schule diese nicht für den Intergruppenvergleich heranziehen, sondern vermehrt auf außerschulische Verhaltensweisen und Eigenschaften und auch Sportlichkeit zurückgreifen, in welchem für Jungen und Männer eine wahrgenommene Überlegenheit der Eigengruppe dominiert und welche damit als geeignet erscheinen, eine positive Distinktheit der Eigengruppe wiederherzustellen.

Folgende Hypothesen sollen im Rahmen dieser Studie überprüft werden: Bei der Einschätzung von Leistungsdimensionen für die Vergabe einer attraktiven Position...

H15 ...bewerten Jungen nach Aktivierung des Stereotyps „Jungen als Bildungsverlierer“ (Experimentalbedingung), relativ zu Jungen der Kontrollgruppe und zu Mädchen beider Bedingungen auf einer Rangskala außerschulische Leistungsdimensionen wichtiger als schulische Leistungsdimensionen.

H16 ...bewerten Jungen der Experimentalgruppe, relativ zu Jungen der Kontrollgruppe und Mädchen beider Bedingungen außerschulische Leistungsdimensionen als *relevanter* für die Vergabe der Position als schulische Leistungsdimensionen.

H17 ... bewerten Jungen der Experimentalgruppe Sportlichkeit als wichtigeres Auswahlkriterium als Jungen der Kontrollgruppe und als Mädchen beider Bedingungen.

Somit wird in der vorliegenden Studie auf einen eher indirekten Mechanismus zur Wiederherstellung positiver Distinktheit fokussiert, bei welchem als soziale Kreativitätsstrategie der relative Status der Eigengruppe anhand der Vergleichskriterien angepasst wird, jedoch die Fremdgruppe kein *direktes* Ziel der Distinktheitsbemühungen darstellt (vgl. *Studie 5*).

10.2. Methode

Teilnehmende. Für diese Untersuchung wurden $N = 126$ Schüler/-innen (51 Jungen und 75 Mädchen) der Klassenstufe 9 einer Berliner integrierten Sekundarschule gewonnen. Das Al-

ter lag durchschnittlich bei $M = 14.73$ Jahren ($SD = .64$, $range = 14 - 16$ Jahre). Auf die Experimentalgruppe entfielen 24 Jungen und 39 Mädchen, in der Kontrollgruppe waren 27 Jungen und 36 Mädchen enthalten. Die Sichtung und initiale Analyse der Testhefte ergab, dass alle Testhefte einen ausreichenden Bearbeitungsstand aufwiesen und somit keine Probanden/-innen von den weiteren Analysen ausgeschlossen werden mussten.

Die Analysen zum Migrationshintergrund der Teilnehmenden zeigen, dass neun von 50 Jungen und sechs von 75 Mädchen angeben, im Ausland geboren zu sein (zur vollen Stichprobengröße fehlende Angaben entsprechen Auslassungen bei der Beantwortung dieser Frage). Mit $\chi^2(1, N = 125) = 2.84$, $p = .09$ weist dieser Unterschied eine marginale Signifikanz in Richtung eines höheren Anteils von Jungen mit direkter Migrationserfahrung auf. Wird die mit den Eltern gesprochene Sprache als zusätzlicher Indikator für eine (indirekte) Migrationserfahrung einbezogen, ergibt sich ein substantiell höherer Anteil mit 21 von 51 Jungen und 31 von 74 Mädchen, welche mit ihren Eltern eine andere Sprache als Deutsch sprechen, wobei hier der Unterschied mit $\chi^2 < 1$, $p = .94$ auf keinerlei bedeutsame Unterschiede zwischen den Geschlechtern hinweist. Betrachtet man die Verteilung der Variablen zum Migrationshintergrund nach Untersuchungsbedingung zeigt sich, dass fünf von 62 Teilnehmenden der Experimentalbedingung und 10 von 63 Probanden/-innen der Kontrollgruppe angeben, im Ausland geboren zu sein. Auf dem Indikator der mit den Eltern gesprochenen Sprache berichten 26 von 62 Teilnehmenden der Experimentalgruppe und 26 von 63 Personen in der Kontrollgruppe, mit den Eltern eine andere Sprache als Deutsch zu sprechen. Auf beiden Indikatoren lassen sich keine signifikanten Unterschiede zwischen Probanden/-innen in der Experimental- und Kontrollgruppe feststellen, $\chi^2 < 1.80$, $p > .18$. Insgesamt bleibt somit festzuhalten, dass insbesondere unter Berücksichtigung des Indikators für indirekte Migrationserfahrung ein hoher Anteil der befragten Schüler/-innen einen Migrationshintergrund aufweist, wobei die Verteilung der Häufigkeiten weder zwischen den Geschlechtern noch in Abhängigkeit der Untersuchungsbedingung systematisch divergiert.

Beim sozioökonomischen Hintergrund, operationalisiert über ausgewählte Fragen zu *home possessions* (Frey et al., 2009), zeigt sich mit einer durchschnittlichen Verfügbarkeit von $M = 5.63$ von 7 abgefragten Gütern ($SD = 1.13$, $range = 2 - 7$) ein vergleichsweise hoher sozioökonomischer Status der befragten Personen. Zwischen den Geschlechtern variiert die Verfügbarkeit der erfragten Güter ebenso nicht systematisch ($M_{\text{Jungen}} = 5.80$, $SD = 1.16$; $M_{\text{Mädchen}} = 5.52$, $SD = 1.11$) wie zwischen den Untersuchungsbedingungen ($M_{\text{EG}} = 5.57$, $SD = 1.13$; $M_{\text{KG}} = 5.69$, $SD = 1.14$), $F < 1.76$, $p > .19$ für alle Haupteffekte und die Interaktion.

Prozedur und Material. Die Teilnehmerinnen und Teilnehmer wurden zu Beginn der Studie randomisiert auf die Experimental- und Kontrollbedingung zugewiesen. Nach einer standardisierten, mündlichen Instruktion zum Umgang mit den Testheften erfolgte die Ausgabe der

Hefte und eine gemeinsame Übung zum Umgang mit den darin enthaltenen, statistischen Informationen sowie die Klärung etwaiger offener Fragen. Zur Erhöhung der Teilnahmemotivation wurde für jedes ernsthaft bearbeitete Testheft eine Belohnung in Höhe von 10 Euro für die Klassenkasse in Aussicht gestellt. Danach begann die Arbeit in den Testheften. In der Experimentalgruppe präsentierten wir dieselbe Statistik zum differenziellen Bildungserfolg von Jungen und Mädchen wie in der vorhergehenden Studie. In der Kontrollgruppe wurde ebenfalls kongruent zur *Studie 5* die Statistik nicht auf Geschlechterunterschiede bezogen, sondern konzentrierte sich vielmehr auf den Wandel des deutschen Schulsystems am exemplarischen Beispiel der Bundesländer Hamburg und Bayern. Zur Sicherstellung einer adäquaten Verarbeitung des präsentierten Datenmaterials wurden wiederum vier Kontrollfragen zum Aufgabenverständnis gestellt. Darauf aufbauend erfolgte die Aufgabenstellung zur Überprüfung der aufgestellten Hypothesen. Zuerst wurde eine einführende Geschichte präsentiert, welche mit den folgenden Aufgaben vertraut macht. Es wurde vorgestellt, dass die Wahlen zur Schulsprecherin/zum Schulsprecher ausstehen und durch eine verstärkte Mitbestimmung von Schüler/-innen bei der Entscheidung für die Wahl geprägt ist. In einem ersten Schritt wurden fünf Bereiche (2 schulische, 2 außerschulische, 1 Sportlichkeit¹¹) durch eine „Vergabekommission“ vorausgewählt, die nun von den befragten Schüler/-innen in eine Rangfolge zu bringen waren (jeder Rang durfte nur einmal vergeben werden). Das individuell am Wichtigsten erachtete Kriterium sollte den oberen Rangplatz einnehmen. Ergänzend wurde eine Liste von schulischen, außerschulischen und sportlichkeitszentrierten Eigenschaften vorgelegt, mit der Bitte diese nach Wichtigkeit des Vorhandenseins der entsprechenden Eigenschaften für die auszuwählende Person von 1 (gar nicht wichtig) bis 10 (sehr wichtig) zu gewichten (in Anlehnung an van Knippenberg & van Oers, 1984). Die vorgelegten Eigenschaften ließen sich dabei wiederum den Dimensionen schulisches Engagement (z.B. Schulnote in Deutsch, Schulnote in Mathematik, gute Mitarbeit in der Schule), außerschulisches Engagement (z.B. in Vereinen aktiv sein, technisch oder handwerklich geschickt sein, in der Freizeit Verantwortung übernehmen) oder Sportlichkeit (z.B. sich im Sport gut auskennen, sportlich sein, sich an Sportwettkämpfen beteiligen) zuordnen. Anschließend wurden die soziodemografischen Variablen wie Alter, Geschlecht, Geburtsort, gesprochene Sprache mit Eltern, Geschwistern, Verwandten und Freunden sowie die Fragen zu den häuslichen Besitztümern zur Bearbeitung präsentiert. Nach vollständiger Bearbeitung der Testhefte wurden die Probanden/-innen vollständig über Inhalt und Ziel der Befragung aufgeklärt und hatten die Möglichkeit, offene Fragen zur Untersuchung ausführlich zu klären. Abschließend erfolgte die symbolische Übergabe der in Aussicht gestellten Entlohnung in Form eines Schecks. Insgesamt betrug die Dauer der Untersuchung ca. 45 Minuten.

¹¹ Sportlichkeit wurde aufgrund der außergewöhnlich hohen Bedeutsamkeit für die männliche Identität (vgl. S. 165f. und die dort zitierten Quellen) nicht in der Kategorie außerschulische Merkmale subsummiert, sondern zur besseren Differenzierbarkeit der Ergebnisse als separate Dimension umgesetzt.

10.3. Ergebnisse

Initiale Analysen. Zunächst wurde erneut überprüft, ob die Kontrollfragen nach Bearbeitung des Treatmentmaterials in ausreichendem Umfang richtig bearbeitet wurden, sodass davon ausgegangen werden kann, dass die zentralen Informationen, welche im präsentierten Material dargeboten wurden durch die Schüler/-innen adäquat verstanden wurden. Für einen Einschluss in die weiterführenden Analysen wurde erneut ein *cut off*-Kriterium von mindestens 50% korrekter Antworten definiert. Mit einem Mittelwert von $M = 71.13\%$ ($SD = 18.13\%$, $range = 50-100\%$) wurde insgesamt ein zufriedenstellendes Niveau an korrekten Antworten auf die Kontrollfragen zum Treatment erreicht. Eine univariate Varianzanalyse mit dem Prozentanteil korrekter Antworten als abhängige Variable und den Faktoren Untersuchungsbedingung und Probanden/-innengeschlecht ergibt einen signifikanten Haupteffekt auf dem Faktor Untersuchungsbedingung, $F(1,122) = 9.39$, $p < .01$, $\eta^2 = .07$ und indiziert damit, dass in der Experimentalbedingung ($M = 75.79\%$, $SD = 18.63\%$) ein deutlich höherer Anteil an Kontrollfragen richtig beantwortet wurde als in der Kontrollgruppe ($M = 66.47\%$, $SD = 16.47\%$). Ebenso ergibt sich ein tendenziell signifikanter Haupteffekt auf dem Faktor Geschlecht, $F(1,122) = 2.82$, $p = .10$, $\eta^2 = .02$, wobei Jungen ($M = 74.02\%$, $SD = 19.18\%$) einen höheren Anteil korrekter Antworten auf die Kontrollfragen aufweisen als Mädchen ($M = 69.17\%$, $SD = 17.23\%$). Die Interaktion zwischen beiden Faktoren verfehlt mit $F < 1$, n.s. das Signifikanzniveau deutlich. Zusammenfassend kann festgehalten werden, dass der Anteil korrekt beantworteter Kontrollfragen zum Treatmentmaterial insgesamt als gut bis zufriedenstellend zu bewerten ist, wobei die für diese Untersuchung besonders relevanten Gruppierungen (Jungen sowie Teilnehmende der Experimentalbedingung) eine höhere Leistung bezüglich der Kontrollfragen aufwiesen – den höchsten Anteil korrekter Antworten erreichten dabei Jungen der Experimentalbedingung mit $M = 79.69\%$, $SD = 19.79\%$.

Ergebnisse zu den Rangdaten. Die von den Schüler/-innen zugeordneten Rangplätze für die fünf vorgelegten schulischen vs. außerschulischen Dimensionen wurden auf Grundlage der besonderen Beschaffenheit dieses Datenmaterials (Ordinalskalierung; abhängige Daten, da jeder Rangplatz nur einmal vergeben werden durfte) einem nichtparametrischen Analyseverfahren nach Kruskal und Wallis (1952) unterzogen. Im Rahmen dieses Verfahrens werden Rangsummen (RS) gebildet, die Anhand einer χ^2 -verteilten Prüfgröße H auf Unterschiede zwischen den untersuchten Gruppen getestet werden können. Niedrigere Ränge (RM , mittlere Ränge) weisen hierbei auf eine größere Bedeutung der abgefragten Merkmale für die Versuchspersonen hin (*Anmerkung:* Der Kruskal-Wallis-Test zieht zur Berechnung der Prüfgröße H dem Skalenniveau der Daten entsprechend Rangsummen heran. Zur besseren Interpretierbarkeit der Daten wird hinsichtlich des unterschiedlichen Anteils von Jungen und Mädchen in der Gesamtstichprobe auf die mittleren Rangwerte zurückgegriffen). Die Analysen zwischen den Geschlechtern machen

deutlich, dass Jungen unabhängig der Untersuchungsbedingung der Kategorie „Beliebtheit bei Anderen“ deutlich mehr Bedeutung für die Vergabe der in der Instruktion vorgestellten, attraktiven Position zuschreiben als Mädchen ($RM_{\text{Jungen}} = 3.20$, $RM_{\text{Mädchen}} = 3.85$, $H = 8.51$, $p < .01$). Gleiches gilt für das Merkmal Sportlichkeit („sportlich sein“, $RM_{\text{Jungen}} = 3.57$, $RM_{\text{Mädchen}} = 4.31$, $H = 5.22$, $p < .05$). Ein umgekehrtes Muster zeigt sich beim schulischen Engagement ($RM_{\text{Jungen}} = 2.59$, $RM_{\text{Mädchen}} = 1.63$, $H = 18.50$, $p < .001$) sowie tendenziell beim Merkmal außerschulisches Engagement ($RM_{\text{Jungen}} = 2.76$, $RM_{\text{Mädchen}} = 2.37$, $H = 3.10$, $p = .08$). Hinsichtlich der schulischen Leistungen kann kein Unterschied bei den zugewiesenen Rängen zwischen Jungen und Mädchen festgehalten werden, beide Gruppen finden diesen Faktor in etwa mittelmäßig wichtig ($RM_{\text{Jungen}} = 3.10$, $RM_{\text{Mädchen}} = 2.88$, $H = .57$, $p = .45$). Zwischen den Untersuchungsbedingungen lassen sich unabhängig vom Geschlecht der Probanden/-inne keine bedeutsamen Effekte auf die Rangplatzzuweisung der fünf abgefragten Merkmale feststellen, alle $H < 1.85$, $p > .17$.

Weiterführende Analysen getrennt nach Geschlecht zeigen, dass zwischen Jungen der Experimental- und Kontrollgruppe lediglich auf dem Merkmal Sportlichkeit ein tendenziell signifikanter Effekt festgestellt werden kann. Jungen der Experimentalgruppe erachten dieses Merkmal mit einem Rang von $RM = 3.08$ tendenziell wichtiger für die Vergabe der attraktiven Position als Jungen der Kontrollgruppe ($RM = 4.00$), $H = 3.02$, $p = .08$. Auf den anderen Merkmalen können keine bedeutsamen Unterschiede bezüglich der Rangwerte festgehalten werden, Beliebtheit: $RM_{\text{EG}} = 3.46$, $RM_{\text{KG}} = 2.96$; außerschulisches Engagement: $RM_{\text{EG}} = 2.92$, $RM_{\text{KG}} = 2.63$; schulische Leistung: $RM_{\text{EG}} = 3.33$, $RM_{\text{KG}} = 2.89$; schulisches Engagement: $RM_{\text{EG}} = 2.63$, $RM_{\text{KG}} = 2.56$; alle $H < 1.90$, $p > .17$). Bei Mädchen zeigen die Testungen keinerlei signifikante Unterschiede zwischen Experimental- und Kontrollgruppe, Beliebtheit: $RM_{\text{EG}} = 3.77$, $RM_{\text{KG}} = 3.94$; außerschulisches Engagement: $RM_{\text{EG}} = 2.49$, $RM_{\text{KG}} = 2.25$; Sportlichkeit: $RM_{\text{EG}} = 4.36$, $RM_{\text{KG}} = 4.25$; schulische Leistung: $RM_{\text{EG}} = 2.82$, $RM_{\text{KG}} = 2.94$; schulisches Engagement: $RM_{\text{EG}} = 1.64$, $RM_{\text{KG}} = 1.61$; alle $H < .97$, $p > .32$.

Die Abbildung 12 fasst die Ergebnisse der Rangdatenanalyse zwischen den Untersuchungsbedingungen sowie getrennt nach Geschlecht zusammen.

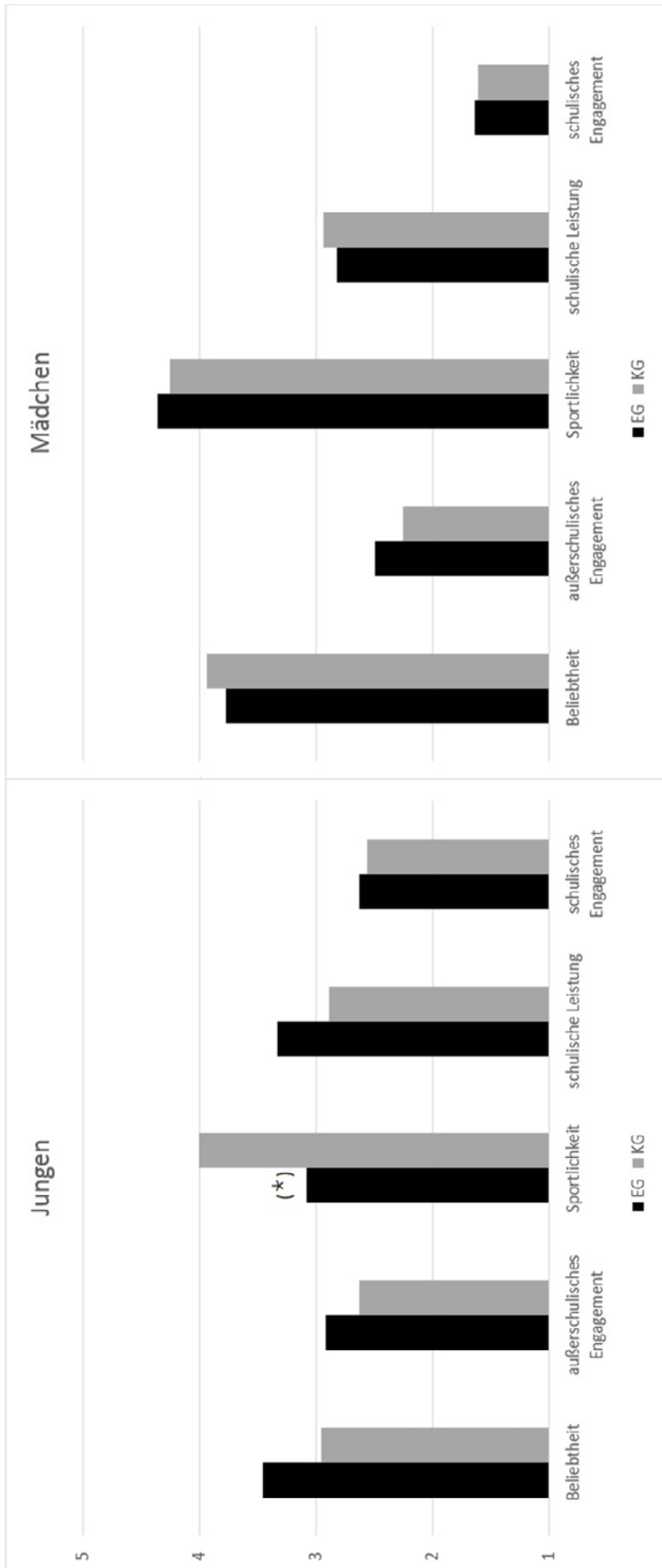


Abbildung 12: Mittlere Ränge je Merkmal nach Geschlecht und Untersuchungsbedingung. Höhere Rangwerte indizieren eine geringere Wichtigkeit der jeweiligen Eigenschaft.

Anm.: Die Prüfstatistik H des zugrundeliegenden Tests zieht zur Berechnung Rangsummen heran. Zur besseren Nachvollziehbarkeit der Ergebnisse wird zur Visualisierung allerdings auf mittlere Rangwerte zurückgegriffen.

(*) = $p \leq .10$

Ergebnisse der geschlossenen Fragen. Im weiteren Verlauf der Untersuchung wurden die Schüler/-innen gebeten, zu vorgegebenen (schulischen, außerschulischen und sportbezogenen) Merkmalen eine Gewichtung (1 = völlig unwichtig, 10 = sehr wichtig) zu vergeben. Die Angaben wurden auf drei Subskalen, schulbezogene Eigenschaften ($\alpha = .76$), außerschulische Eigenschaften ($\alpha = .61$) und Sportlichkeit ($\alpha = .86$) als Mittelwerte zusammengefasst und einer 2 (Geschlecht der Versuchsperson) \times 2 (Untersuchungsbedingung) \times 3 (Subskalen: schulisch, außerschulisch, sportbezogen) *mixed model*-ANOVA mit den drei o.g. Subskalen als Messwiederholungsfaktor unterzogen. Die Ergebnisse zeigten einen Haupteffekt auf dem Messwiederholungsfaktor, welcher darauf hindeutete, dass der Teilkomponente schulischer Merkmale unabhängig vom Geschlecht der befragten Person und der Untersuchungsbedingung die größte Relevanz zugeschrieben wurde ($M = 7.31$, $SD = 1.14$), gefolgt von außerschulischen Eigenschaften ($M = 5.35$, $SD = 1.45$) und schlussendlich Sportlichkeit ($M = 5.04$, $SD = 2.32$). Weiterhin ergab sich ein Haupteffekt des Versuchspersonengeschlechts, $F(1,122) = 5.26$, $p < .05$, $\eta^2 = .04$, in Richtung einer höheren eingeschätzten Relevanz aller abgefragten Bereiche bei Jungen ($M_{\text{Schule}} = 7.44$, $SD = 1.20$; $M_{\text{außerschulisch}} = 5.67$, $SD = 1.61$; $M_{\text{Sport}} = 5.65$, $SD = 2.64$) im Vergleich zu Mädchen ($M_{\text{Schule}} = 7.22$, $SD = 1.10$; $M_{\text{außerschulisch}} = 5.13$, $SD = 1.30$; $M_{\text{Sport}} = 4.64$, $SD = 1.99$). Weitere Haupteffekte fehlten mit $F < 1$, n.s. die statistische Bedeutsamkeit.

Die berichteten signifikanten Haupteffekte wurden durch eine marginal signifikante Interaktion aus dem Messwiederholungsfaktor und dem Faktor Geschlecht ergänzt, $F(2,121) = 2.40$, $p = .09$, $\eta^2 = .04$. Wie aus den soeben dargestellten deskriptiven Werten deutlich wurde, fiel der Unterschied zwischen den Einschätzungen von Jungen und Mädchen hinsichtlich der Sportlichkeit am größten aus, deren Bedeutsamkeit ein signifikanter post hoc-Test bestätigte, $t(124) = -2.32$, $p < .05$, $d = -.43$, wobei Jungen der Experimentalbedingung Sport als relevanter einschätzten, als Mädchen der Kontrollgruppe, $t(58) = 2.22$, $p < .05$, $d = .58$, oder tendenziell Mädchen der Experimentalgruppe, $t(61) = -1.80$, $p = .08$, $d = -.45$. Ebenfalls wurde der Unterschied auf der Komponente außerschulische Eigenschaften des Messwiederholungsfaktors im post hoc-Test statistisch bedeutsam, $t(124) = -2.00$, $p = .05$, $d = -.37$, wohingegen bezüglich der schulischen Eigenschaften keine signifikanten Unterschiede zwischen Jungen und Mädchen festgehalten werden konnten, $t = -1.04$, $p = .30$. Weitere Interaktionen konnten nicht festgestellt werden, $F < 1$, n.s. Die Dreifachinteraktion aus allen in die Varianzanalyse eingegangenen Variablen wies mit $F < 1$ ebenfalls keine statistische Bedeutsamkeit auf und indizierte auch unter Rückgriff auf die bisher berichteten Ergebnisse insbesondere keinen systematischen Einfluss der Untersuchungsbedingung auf die Relevanzeinschätzung der befragten Schülerinnen und Schüler.

Abbildung 13 stellt die durchschnittlichen Relevanzeinschätzungen nach Geschlecht und Untersuchungsbedingung zusammenfassend grafisch dar.

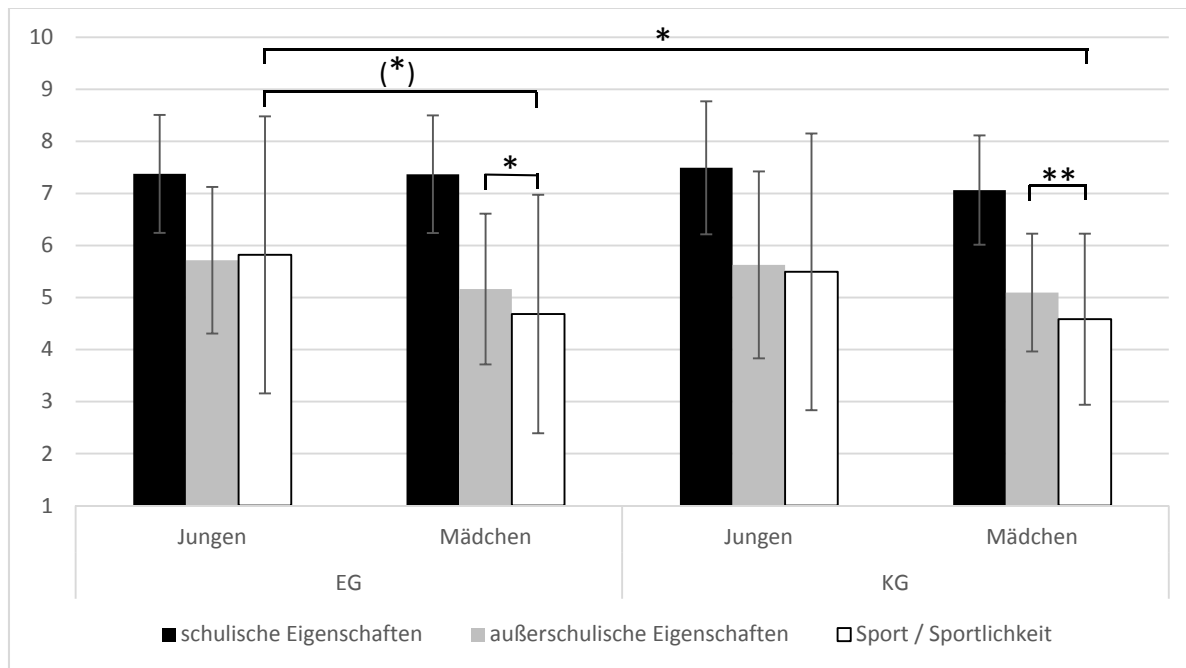


Abbildung 13: Einschätzung der Relevanz von schulischen und außerschulischen Eigenschaften für die Vergabe einer attraktiven Position nach Geschlecht und Experimentalgruppe.

Anm.: Die Unterschiede zwischen schulischen Eigenschaften (schwarzer Balken) und den zwei anderen abgefragten Subskalen (grauer und weißer Balken) sind stets mit $p \leq .02$ signifikant. Weitere signifikante Unterschiede im post hoc-Testungen sind in der Abbildung enthalten, (*) = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Zusammenfassend lässt sich festhalten, dass Jungen zwar außerschulischen und sportbezogenen Eigenschaften mehr Gewicht für die Vergabe einer attraktiven Position beigemessen haben, dieser Effekt aber (mit Ausnahme der Dimension Sportlichkeit) unabhängig von der Experimentalbedingung auftrat. Ebenso unbeeinflusst durch das experimentelle Treatment erachteten Mädchen Sport als vergleichsweise unwichtigstes Kriterium, wohingegen Jungen nicht hinsichtlich der Relevanz außerschulischer und sportbezogener Merkmale differenzierten. Zusätzlich bleibt allerdings zu berücksichtigen, dass sowohl Jungen als auch Mädchen beider Untersuchungsbedingungen schulisches Engagement einhergehend mit einer hohen schulischen Leistungsfähigkeit als relevantestes Kriterium ansahen.

10.4. Diskussion

Im Rahmen dieser Studie wurde eine weitere Strategie aus der Theorie der sozialen Identität (Tajfel & Turner, 1986) zur Wiederherstellung positiver Distinktheit als Reaktion auf die Aktivierung des schulversagenden Jungen überprüft. Insgesamt zeigt sich, dass bei der Aufgabe zur Rangsortierung ausschließlich Effekte zwischen Jungen und Mädchen zu beobachten sind, welche nicht systematisch durch das experimentelle Treatment beeinflusst werden. Jungen nominieren hierbei unabhängig von der experimentellen Bedingung die Beliebtheit bei anderen Schüler/-innen, Sportlichkeit und tendenziell auch außerschulisches Engagement auf niedrigere

Rangplätze, resp. messen diesen Merkmalen eine erhöhte Bedeutsamkeit bei. Schulischem Engagement wird zusätzlich von Jungen ein signifikant höherer Rangplatz, resp. eine geringere Bedeutsamkeit zugeschrieben. Ein differenzieller Einfluss der Untersuchungsbedingung kommt lediglich tendenziell auf der Dimension Sportlichkeit zum Tragen, wo Jungen der Experimentalgruppe einen etwas geringeren Rangplatz relativ zu Jungen der Kontrollbedingung vergeben. Ergänzt um die geschlossenen Fragen zur Gewichtung verschiedener Eigenschaftsbereiche zeigt sich, dass unabhängig von Geschlecht oder Untersuchungsbedingung schulischen Eigenschaften die höchste Relevanz zugeschrieben wird, gefolgt von außerschulischen Merkmalen und Sportlichkeit. Weiterhin werden alle drei abgefragten Bereiche von Jungen als relevanter eingeschätzt, als von Mädchen. Einflüsse des Geschlechts zeigen sich nur tendenziell dahingehend, dass Jungen der Experimentalgruppe der Dimension Sportlichkeit eine etwas höhere Relevanz für die Vergabe einer attraktiven Position zusprechen, als Mädchen beider Untersuchungsbedingungen. Insbesondere zum Einfluss des experimentellen Treatments muss festgehalten werden, dass hier nur wenige Effekte auftreten, welche dann zumeist lediglich eine tendenzielle Signifikanz erreichen.

Bezugnehmend auf die Hypothesen bleibt somit festzuhalten, dass die Hypothese 15 abzulehnen ist. Jungen wählen zwar außerschulische Leistungsdimensionen incl. Sportlichkeit auf einem geringeren Rangplatz und erachten diese Dimensionen somit als ein bedeutsameres Auswahlkriterium als Mädchen, allerdings konnte kein differenzieller Einfluss des experimentellen Treatments festgestellt werden – die Rangplatzvergabe folgte hier für Jungen *generell* dem o.g. Muster. Hypothese 16 kann ebenfalls als nicht bestätigt angesehen werden, da Jungen entgegen der Erwartung in beiden Untersuchungsbedingungen nicht systematisch hinsichtlich der Einschätzung der Relevanz von schulischen, außerschulischen und sportbezogenen Eigenschaften und Verhaltensweisen für die Vergabeentscheidung divergieren. Da Jungen unabhängig von der Experimentalbedingung der Kategorie Sport und Sportlichkeit einen niedrigeren Rangplatz und eine tendenziell höhere Relevanz zugeschrieben haben als Mädchen und wiederum kein Einfluss der Experimentalbedingung festgestellt werden kann, bleibt auch Hypothese 17 abzulehnen. Schlussendlich kann nur festgehalten werden, dass bei Mädchen sowohl bei der Vergabe der Rangplätze als auch bei der Relevanzeinschätzung keine Unterschiede in Abhängigkeit der Zugehörigkeit zur Experimental- oder Kontrollgruppe gezeigt werden kann.

Es erscheint im Hinblick auf die ausbleibenden Befunde auf dem Faktor Untersuchungsbedingung neben methodischen Problemen (vgl. Limitationen) denkbar, dass die Vergabe der Position einer Schulsprecherin/eines Schulsprechers auf Grund der generell weiblichen Konnotation des Kontextes Schule (Hannover & Kessels, 2011) unter Umständen zu einer automatischen Aktivierung von Stereotypen (Devine, 1989) geführt hat und damit auch in der Kontrollgruppe entweder darin resultiert, dass die zu vergebende Position eher mit schulischen (weil

weiblichen) Eigenschaften assoziiert wird, oder aber bei Jungen der Kontrollgruppe eine nicht intendierte Stereotypenbestätigung (McNatt, 2000) ausgelöst wurde und damit einhergehend eine ähnliche Charakterisierung der idealen Person für die zu vergebende Position wie in der Experimentalgruppe getriggert wird. Es bleibt also zu diskutieren, inwieweit durch die mit der zu vergebenden Position verbundenen Merkmale evtl. zu einer automatischen Stereotypenaktivierung geführt haben, welche Effekte durch das experimentelle Treatment überlagern. Betten-court, Charlton, Dorr & Hume (2001) zeigen metaanalytisch, dass die Eigengruppenfavorisierung (auch auf irrelevanten Dimensionen im Sinne sozialer Kreativität) bei Gruppen mit hohem sozialen Status immer stark ausgeprägt ist, wohingegen sich sozial niedriger gestellte Gruppen nur dann auf irrelevante Vergleichsdimensionen beziehen, wenn die Gruppengrenzen undurchlässig sind. Trotz dessen, dass die Kontrollgruppe keine expliziten Bedrohung der sozialen Identität ausgesetzt wurde, kann durch den Kontext der zu vergebenden, attraktiven Position auf einen schulischen Kontext geschlossen werden und in Verbindung mit der Impermeabilität der Gruppengrenzen, welche sich *qua* Geschlecht definieren und damit einen Gruppenwechsel unmöglich machen, nichtsdestotrotz soziale Kreativität im Sinne einer stärkeren Fokussierung auf irrelevante Vergleichsdimensionen (Ellemers & Van Rijswijk, 1997; Ellemers et al., 1993) auch bei Jungen der Kontrollgruppe ausgelöst haben.

Weiterhin kann vermutet werden, dass Geschlecht im Jugendalter eine sehr zentrale Dimension darstellt, welche über verschiedene Kontexte in diesem Entwicklungsstadium hinweg hochgradig salient ausfällt (Basow & Rubin, 1999; Galambos, Almeida & Peterson, 1990) und zusätzlich Jungen die Mitgliedschaft in Gruppen als zentraleres Merkmal des Selbst erachten (Cross & Madson, 1997) sowie eine stärkere kollektive Interdependenz aufweisen, wohingegen Mädchen und Frauen eher als relational interdependent beschrieben werden (Gabriel & Gardner, 1999). Möglicherweise stellt das Geschlecht und die damit assoziierten Eigenschaften damit auch ohne Hinzunahme des experimentellen Treatments bereits eine derartig zentrale Kategorie dar, dass auch Jungen der Kontrollgruppe ein bestimmtes Maß an (schulbezogenen) Geschlechterdisparitäten wahrnehmen und dementsprechend agieren. Es besteht somit die Möglichkeit, dass Geschlecht als zentrale Dimension im Altersbereich der untersuchten Stichprobe eine wichtige Rolle im Rahmen der sozialen Wahrnehmung spielt und damit einhergehend das applizierte experimentelle Treatment nicht ausreichend war, um diesen grundlegenden Effekt substanziell zu verstärken.

Der Kategorie der Sportlichkeit kommt bei den vorliegenden Befunden offenbar eine besondere Bedeutung für Jungen zu, da hier sowohl hinsichtlich der Rangdaten als auch bei der Einschätzung der Relevanz sportbezogener Eigenschaften konsistent eine höhere Bedeutsamkeit dieser Komponente für Jungen festgestellt werden konnte. Sport und Sportlichkeit scheint hier eine Dimension zu sein, welche noch als besonders maskulin konnotiert wahrgenommen

wird und sich daher als Vergleichsdimension im Sinne einer Strategie der sozialen Kreativität als besonders verfügbar und geeignet darstellt. Verschiedene Studien konnten zeigen, dass Sport als maskulines Merkmal einen hohen Stellenwert im Rahmen der männlichen Identität einnimmt und daraus resultiert, dass Jungen und Männer sich stärker mit dieser Thematik identifizieren, auseinandersetzen und schlussendlich auch eine stärkere Involviertheit in entsprechende Tätigkeiten berichten (Bhana, 2008; Ricciardelli et al., 2006; Mutz, 2009; Schmalz & Kerstetter, 2006; Larson, 2000; Eccles & Harold, 1991). Möglicherweise kann zur Erklärung der fehlenden Effekte hinsichtlich der Untersuchungsbedingung beitragen, dass die Relevanz von Sport ferner im Sinne einer positiven Bewertung des eigenen Körpers und der Betonung von Körperlichkeit für die männliche Identitätsbildung im Jugendalter eine derartig wichtige Dimension darstellt (Schaufeler, 2002; Seiffge-Krenke, 1994; Harter, 1990), dass die experimentelle Beeinflussung durch ein negatives Gruppenstereotyp („Jungen als Bildungsverlierer“) nicht ausreicht, die ohnehin hohe Relevanz dieses Konstruktes maßgeblich zu intensivieren. Gegen diese Interpretation spricht, dass alle anderen abgefragten Bereiche gleichwertig wichtig oder sogar wichtiger eingeschätzt werden. Dies legt nahe, dass durchaus ein Spielraum zur Erhöhung der Betonung von Sport und Körperlichkeit gegeben ist, welcher sowohl theoretisch fundiert (Tajfel & Turner, 1986) als auch empirisch erwartbar (Jackson et al., 1996) als Reaktion auf eine Bedrohung der sozialen Identität genutzt werden sollte.

Zusammenfassend stehen die ausbleibenden Effekte in der vorliegenden Studie in Kontrast zu zahlreichen publizierten Arbeiten zur Eigengruppenfavorisierung mittels sozialer Kreativität, welche zum Großteil durch die Theorie vorhergesagte Effekte (Eigengruppenbevorzugung der niedrigeren Statusgruppe auf für den Vergleich irrelevanten Dimensionen) aufzeigen konnten (u.a. Ashford & Kreiner, 1999; Ellemers & van Rijswijk, 1997; Elsbach & Kramer, 1996; Jackson et al., 1996; Terry & Callan, 1998; zusammenfassend: Mullen et al., 1992). Diese Studien fokussieren allerdings zumeist auf das Erwachsenenalter oder den organisationalen Kontext und beziehen damit die Besonderheiten der in dieser Studie untersuchten Altersspanne, dem Jugendalter, nicht konsequent mit ein. Die Phase des Jugendalters charakterisiert sich dabei durch die zentrale Thematik der Identitätsentwicklung, d.h. das Selbst differenziert sich im Jugendalter zunehmend aus (Harter, 2003; Harter, Bresnick, Bouchey & Whitesell, 1997) und wird im Vergleich zum frühen Erwachsenenalter als weniger ausdifferenziert und stabil wahrgenommen (Kernis & Goldman, 2003; Labouvie-Vief, 2003). Es erscheint daher möglich, dass u.U. weniger Dimensionen für die Anwendung sozialer Kreativitätsstrategien zur Wiederherstellung positiver Distinktheit im Jugendalter verfügbar sind. Auch bleiben (Geschlechter-) Unterschiede hinsichtlich des Bewältigungsverhaltens einer sozialen Bedrohung (wie in dieser Studie appliziert) zu berücksichtigen: Eschenbeck, Kohlmann & Lohaus (2007) weisen darauf hin, dass Mädchen in der Adoleszenz zumeist soziale Unterstützung suchen und ein aktiveres Problemlöseverhalten

als Reaktion auf einen sozialen Stressor zeigen, wohingegen Jungen eher vermeidendes Verhalten präferieren (vgl. auch Causey & Dubow, 1992), was möglicherweise mit der Art der Erfassung der abhängigen Variable in der vorliegenden Studie konfligierte. Eine genauere Prüfung dieser theoretischen Argumentation in weiterer empirischer Forschungsarbeit erscheint allerdings zwingend notwendig, um die Richtigkeit der getroffenen Vermutungen zu überprüfen.

Limitationen. Es konnte im Hinblick auf die Operationalisierung der abhängigen Variable nicht umfassend sichergestellt werden, ob und inwiefern die dargebotene, attraktive Position tatsächlich hinreichend attraktiv wahrgenommen wird und mit welchen Wahrnehmungen diese Position einhergeht. Zwar ist anzunehmen, dass die Position sowohl für Jungen (Führung, Prestige und Macht als sehr maskuline Komponenten; Koenig, Eagly Mitchell & Ristikari, 2011) als auch für Mädchen (schulisches Engagement, Strebsamkeit, sozialer Nutzen als feminine Komponenten) ansprechend wahrgenommen wird. Nichtsdestotrotz kann nicht mit Sicherheit festgestellt werden, ob die Position ausreichend attraktiv war, um soziale Kreativität auszulösen – wenngleich Schulsprecher/-innen positiv wahrgenommen werden (Grundmann & Kramer, 2001) und positive *Outcomes* aus einer entsprechenden Position zu erwarten sind (Schreiter & Bürger, 2006). Davon unberücksichtigt sollte bei einer Replikation der Befunde eher eine attraktive Position gewählt werden, die durch sorgfältige Vortestungen nachgewiesenermaßen von beiden Geschlechtern als attraktiv wahrgenommen wird. Zusätzlich bleibt fraglich, ob die zur Messung der sozialen Kreativität angebotenen Dimensionen und Verhaltenseigenschaften ideal operationalisiert wurden. Es kann anhand der vorliegenden Daten nicht sicher festgestellt werden, inwiefern die dargebotenen Dimensionen (zur Rangreihenbildung) bzw. Eigenschaften für Jungen neben der in der Einführung beschriebenen theoretischen Relevanz tatsächlich auch praktisch eine ernstzunehmende Alternative zu schulischem Erfolg darstellen – insbesondere unter Berücksichtigung, dass die zu vergebende, attraktive Position einem schulischen Kontext entstammt. Die im Rahmen dieser Studie zu vergebende, attraktive Position wurde zur Erhöhung des Bezugs zur Lebenswelt der untersuchten Schüler/-innen auf einen schulischen Kontext bezogen. *Studie 1* dieser Arbeit weist allerdings darauf hin, dass Jungen auch ohne explizite Aktivierung eines Stereotyps spontan bewusst ist, dass ihre Eigengruppe als schulisch weniger engagiert und erfolgreich wahrgenommen wird. Die Verfügbarkeit dieses Wissens vorausgesetzt kann somit auch ohne explizites Ansprechen auf das Stereotyp schulversagender Jungen eine soziale Bedrohungssituation entstanden sein und entsprechende Reaktionen nach sich gezogen haben. Aus den vorliegenden Daten kann nicht sicher geschlossen werden, ob und in welchem Ausmaß dies der Fall ist, was für zukünftige Studien eine zusätzliche Erhebung aktueller emotionaler Zustände (wie Angst, Erregung, etc.; z.B. Statuskomponente des *state trait anxiety inventory*; Laux, Glanzmann, Schaffner & Spielberger, 1981) empfehlenswert macht, um diese Schlussfolgerung zu ermöglichen. Es könnte sich darüber hinaus nichtsdestotrotz als günstig

erweisen, u.U. eine Position in einem neutraleren Kontext vergeben zu lassen und zusätzlich sicherzustellen, dass angebotene Dimensionen zum Wechsel des sozialen Vergleichs als mögliche und ernstzunehmende Alternative angesehen werden. Ferner kann in Betracht gezogen werden, gar keine alternativen Vergleichsdimensionen anzubieten, sondern vielmehr nur die unterscheidende Variable (schulischer Erfolg) hinsichtlich deren Erwünschtheit als Kriterium für die Vergabeentscheidung heranzuziehen (ähnlich Jackson et al. 1996) und ggf. als Reaktion auf die o.g. Probleme bei der geschlossenen Erfassung zusätzlich um offene Kommentare („*welche Kriterien sollten zusätzlich zur Entscheidung herangezogen werden?*“) zu ergänzen, welche dann klassifiziert und einer Clusteranalyse unterzogen werden können. Es kann daher zusammenfassend nicht ausgeschlossen werden, dass die soeben beschriebene Problematik die im Ergebnisteil dargestellten Befundmuster zusätzlich beeinflusst hat.

Abschließend bleibt hinsichtlich der Testkonstruktion kritisch anzumerken, dass erneut (wie in *Studie 5*) Unterschiede in den korrekten beantworteten Kontrollfragen bezüglich des Treatmentmaterials zwischen Experimental- und Kontrollgruppe festgestellt wurden. Wenngleich die interessierende Gruppe (Jungen der Experimentalbedingung) das höchste Maß korrekter Antworten erreichen konnte und demnach diese Unterschiede für die Beantwortung der Fragestellung wenig problematisch erscheint, sollte durch eine kritische Überprüfung der Aufgabenschwierigkeit und ggf. Anpassungen bezüglich der Kontrollfragen eine bessere Vergleichbarkeit zwischen den experimentellen Bedingungen hergestellt werden. Ferner weist die Reliabilität der eingesetzten Skala „*außerschulische Merkmale*“ gerade noch ein ausreichendes Zuverlässigkeitsniveau auf (trotz des Umstandes, dass mit dieser Skala ein vergleichsweise breites Spektrum an Eigenschaften erfasst wird), wobei auch durch Entfernen einzelner Items aus der Gesamtskala keine substanzielle Erhöhung der Reliabilität erreicht werden kann. Dies indiziert, dass eine Neukonstruktion der Skala erwägt werden sollte.

Zu potentiell moderierenden oder mediierenden Variablen, wie z.B. der *gender identification* oder wahrgenommenen Durchlässigkeit der Gruppengrenzen (Jackson et al., 1996) gilt wie in der Diskussion zur vorherigen Studie festgestellt auch für diese Studie, dass ein konsequenterer Einbezug derartiger Variablen in zukünftigen Arbeiten wünschenswert wäre.

11. Studie 8: Wiederherstellung positiver Distinktheit III

Wechsel sozialer Vergleichsgruppen als Bewältigungsstrategie auf das Stereotyp des männlichen Schulversagers.

11.1. Überblick und Fragestellung

Im Folgenden soll eine weitere Strategie zur Wiederherstellung positiver Distinktheit überprüft werden, welche aus der Theorie der sozialen Identität (Tajfel & Turner, 1986) abgeleitet werden kann. Es wird postuliert, dass als Reaktion auf eine Bedrohung der sozialen Identität die Vergleichsgruppe dahingehend geändert werden kann, dass selbstwerterhaltende bzw. -erhöhende Abwärtsvergleiche angestellt werden, sodass das bedrohte Individuum im Vergleichsprozess besser abschneidet, als beim Vergleich mit der ursprünglichen Fremdgruppe. Diese Annahmen werden durch die Theorie des sozialen Vergleichs (Festinger, 1954) unterstützt. Menschliches Erleben und Verhalten wird grundlegend durch soziale Vergleiche mitbestimmt, um eigenes Handeln in spezifischen (sozialen) Situationen zu reflektieren und zu bewerten sowie um die Verhältnismäßigkeit eigenen Verhaltens in mit geltenden sozialen Normen abgleichen zu können. Dem sozialen Vergleich kommt somit sowohl für intra- als auch für interpersonelles Verhalten eine große Bedeutung zu, da diese sowohl Auswirkungen auf die Selbstbewertung als auch auf zwischenmenschliches Verhalten haben können (Kahneman & Miller, 1986; Mussweiler, 2003). Wie genau soziale Vergleiche auf Verhalten und Erleben wirken ist maßgeblich durch die Richtung des Vergleichs determiniert. Genauer geht Wills (1981) davon aus, dass Vergleiche mit einer statusniedrigeren Gruppe selbstwerterhöhend wirken, wohingegen Aufwärtsvergleiche eine motivierende und leistungssteigernde Wirkung aufweisen (siehe auch Taylor & Lobel, 1989). Abwärtsvergleiche sind demnach durch negativen Affekt im Rahmen eines sozialen Vergleichs ausgelöste Prozesse, durch welche das bedrohte Selbst und ein beeinträchtigt Wohlbefinden durch den Vergleich mit weniger erfolgreichen Anderen wiederhergestellt werden kann. Solche Vergleichsprozesse können der Theorie folgend sowohl aktiv erfolgen, in dem der relative Status einer Vergleichsgruppe abgewertet oder durch aktives (hostiles) Handeln vermindert wird (vgl. *Studie 5*) oder aber auf eine passive Art und Weise, indem eine natürlich vorhandene, schlechter gestellte Fremdgruppe für den sozialen Vergleich herangezogen wird (Suls & Wills, 1991; Suls & Wheeler, 2000). Anzumerken bleibt, dass trotz des selbstwerterhöhenden Potentials von Abwärtsvergleichen eine Ambivalenz der Gefühle postuliert wird, d.h. abwärtsgerichtete Vergleiche sind nicht ausschließlich mit einer positiven Emotionalität besetzt, da diese das Bedürfnis nach einem gerechten Vergleich verletzen. Zusätzlich werden abwärtsgerichtete Vergleiche wahrscheinlicher, wenn ein niedrig ausgeprägter Selbstwert vorliegt (Wills, 1981) sowie wenn die Vergleichsdimension für das Individuum relevant ist und als Ursache für den Statusunterschied stabile Faktoren zugeschrieben werden (Frey, Dauheimer, Parge & Haisch, 2001). Auf weitere Einflussfaktoren für das Zustandekommen einer

Selektion eines niedrigeren Vergleichsstandards wie einer Bedrohung der sozialen Identität (wie in dieser Studie experimentell induziert), Selbstwert und Attributionsstil weisen Taylor, Wayment & Carrillo (1996) hin. Ebenso kommt der medialen Darstellung für das Zustandekommen von sozialen Vergleichen eine besondere Bedeutung zu (Jones, 2001), wobei der verstärkte Einfluss der Medien insbesondere bei Frauen und vor allem bei körperbezogenen, nicht-kognitiven Faktoren nachgewiesen wurde (z.B. Strahan, Wilson, Cressmann & Buote, 2006).

Es kann festgehalten werden, dass einerseits der Umgang mit einer Bedrohung der sozialen Identität durch einen sozialen Vergleich mit einer schlechter abschneidenden Fremdgruppe dazu führt, die noch bedrohlichere Wirkung eines Aufwärtsvergleichs abzuwenden (Tesser, Millar & Moore, 1988) und somit das eigene Selbst zu schützen (Gibbons, 1986; Wills, 1981), wenngleich eine gewisse emotionale Ambivalenz bestehen bleibt und bei kurzfristiger Exposition einer Bedrohung ein Abwärtsvergleich weniger wahrscheinlich als emotional negativ wahrgenommen wird, als bei andauernden Bedrohungssituationen (Gibbons & Gerrard, 1991; Taylor et al., 1996; Taylor, Aspinwall, Giuliano & Reardon, 1993). Ferner ist abzuleiten, dass abwärtsgerichtete Vergleiche zwar den Selbstwert erhalten oder erhöhen, jedoch motivational keine Veränderungsprozesse anstoßen und somit zu einer Verfestigung von Statusunterschieden beitragen können. Im Rahmen dieser Studie soll dieser passive Mechanismus im Kontext des Paradigmas „Jungen als Bildungsverlierer“ untersucht werden. Angewendet auf diese Problemstellung in Verbindung mit den soeben ausgeführten theoretischen und empirischen Grundlagen ist zu erwarten, dass Jungen als Reaktion auf die Konfrontation mit dem selbstwertbedrohenden Stereotyp ihres Schulversagens soziale Vergleichsprozesse in Richtung abwärtsgerichteter Vergleiche anwenden, um a) ihr Selbstbild zu schützen und darüber hinaus b) wieder eine positive Distinktheit von der besser abschneidenden Gruppe der Mädchen zu erreichen. Empirisch soll dies unter Rückbezug auf die Annahmen der Theorie der sozialen Identität in Ergänzung um die Theorie der sozialen Vergleichsprozesse anhand folgender Hypothesen überprüft werden:

H18 Jungen der Experimentalgruppe wählen im Vergleich zu Jungen der Kontrollgruppe eher Vergleichsgruppen, die auf der relevanten Vergleichsdimension (Schulleistung) unterlegen sind.

H19 Jungen und Mädchen wählen unabhängig von der Experimentalbedingung bei der Wahl der Vergleichsgruppe häufiger das eigene Geschlecht zur Heranziehung sozialer Vergleiche.

Für das Zustandekommen von Aufwärtsvergleichen, d.h. einem Vergleich mit einer höheren Schulform, soll explorativ und ohne Formulierung einer expliziten Hypothese geprüft werden, inwiefern dieser durch das experimentelle Treatment begünstigt wird. Denkbar wäre hierbei neben einer erhöhten Bereitschaft von aufwärtsgerichteten Vergleichen bei Jungen der Kon-

trollgruppe auch eine verstärkte Neigung zu Aufwärtsvergleichen bei Mädchen der Experimentalbedingung in Reaktion auf die Darstellung der Überlegenheit der eigenen Gruppe.

11.2. Methode

Teilnehmende. Bei der Stichprobe zu dieser Studie handelt es sich um die $N = 87$ Schüler/-innen (41 Jungen und 46 Mädchen), welche bereits für *Studie 3* akquiriert und ausgewertet wurden. Die genaue Beschreibung dieser Stichprobe wurde bereits im Abschnitt 7.2 (Seite 118) vorgenommen und wird daher an dieser Stelle nicht nochmals detailliert wiederholt.

Prozedur und Material. Wie bei der Prozedur zu *Studie 3* ausführlicher beschrieben fand die Untersuchung in zwei aufeinanderfolgenden Schulstunden in einer 9. Klasse eines Berliner Gymnasiums statt. Die Schüler/-innen wurden mittels einer standardisierten Instruktion begrüßt und mit dem Ablauf der Testsituation vertraut gemacht. Darauf erfolgte die Ausgabe der Testhefte, wobei randomisiert hälftig ein Testheft der Experimentalbedingung sowie der Kontrollbedingung ausgegeben wurde. Vor Beginn der Arbeit mit den Testheften wurden die darin enthaltenen Aufgaben- und Antwortformate erläutert sowie der Umgang mit den Leistungsaufgaben sowie Antwortformaten bei Fragebögen erklärt. Nach der Möglichkeit Rückfragen zu stellen begann die eigentliche Arbeit mit dem Testheft, wobei zunächst in randomisierter Abfolge 15 Aufgaben zum Leseverständnis (Deutsch) und 12 Aufgaben zur Mathematik mit teilweise offenem und teilweise geschlossenem Antwortformat zu bearbeiten waren (entnommen VERA 8; Institut zur Qualitätsentwicklung im Bildungswesen, 2008). Nach Bearbeitung der Leistungsaufgaben wurde das experimentelle Treatment appliziert, welches aus der Bearbeitung einer einfachen, kurzen Statistik bestand (Tabellen und Abbildungen) und in der Experimentalgruppe Leistungsunterschiede in der Schule und beim Erwerb schulischer Zertifikate deutlich machte (Statistisches Bundesamt, 2014a) bzw. in der Kontrollgruppe ohne explizite Erwähnung von Geschlecht die Verteilung von Schüler/-innen auf die verschiedenen Schulformen in zwei Bundesländern zum Gegenstand hatte. Im Anschluss an die für *Studie 3* notwendige Angabe, welche Aufgaben im weiteren Verlauf der Untersuchung zur Bearbeitung vorgelegt werden sollen (Erfassung der Zielorientierung) erfolgte der mündliche Hinweis, dass für weitere Aufgaben leider kein ausreichender zeitlicher Rahmen mehr zur Verfügung steht. Abschließend wurde den Schüler/-innen angekündigt, dass Sie nun vielmehr ihre Leistungsaufgaben ansehen können und ihre eigene Leistung der von verschiedenen Vergleichsgruppen gegenüberstellen können. Hierzu sollte eine Rangreihe gebildet werden, welche Vergleichsgruppe für den sozialen Vergleich am meisten bis am wenigsten präferiert wird. Es gab die Möglichkeit eines parallelen Vergleichs mit der eigenen Schulform (Gymnasium), die Möglichkeit eines Aufwärtsvergleichs (Studienanfänger/-innen), sowie zwei mögliche Abwärtsvergleiche (Realschule, Hauptschule). Zusätzlich

konnte für jede der soeben genannten Vergleichsgruppe das Geschlecht angegeben werden (z.B. Jungen, Gymnasium; Mädchen, Gymnasium usw.). Jeder Rang durfte nur einmal vergeben werden, es resultierte somit eine Rangreihe von 1 (am meisten präferierte Vergleichsgruppe) bis 8 (am wenigsten präferierte Vergleichsgruppe). Nach Vergabe der Rangplätze erfolgte die Abfrage einiger soziodemografischen Daten wie Alter, Geschlecht, Migrationshintergrund und sozioökonomischer Status gefolgt von einer ausführlichen Aufklärung der Schüler/-innen über Inhalte und Ziele der Untersuchung. Die Gesamtdauer der Untersuchung betrug ca. 60 Minuten.

11.3. Ergebnisse

Initiale Analysen. Die initialen Analysen zu den Fragen zur Manipulationsüberprüfung entsprechen den für Studie 3 dargestellten Kennwerten (ab S. 125) und werden daher an dieser Stelle nicht nochmals detailliert wiedergegeben. Es galt wie für alle bisher dargestellten Studien ein *cut off*-Kriterium in Höhe von mindestens 50% korrekter Antworten im *manipulation check* für den Einschluss in die weiterführenden Analysen. Im Mittel wurden $M = 65.23\%$ ($SD = 14.16\%$, $range = 50 - 100\%$) der Fragen zur Manipulationsüberprüfung richtig beantwortet, es finden sich keinerlei bedeutsamen Unterschiede zwischen den Untersuchungsbedingungen, den Geschlechtern oder auf dem Interaktionsterm beider Faktoren.

Analyse der Rangdaten. Das Skalenniveau der für die Untersuchung dieser Fragestellung eingesetzten Materialien (Ordinalskalenniveau) schränkte die Analysemöglichkeiten ein, weshalb unter methodologischen Gesichtspunkten zunächst auf eine binär-logistische Regressionsanalyse zurückgegriffen wurde, deren Ergebnisse um verschiedene nichtparametrische Verfahren (Chi-Quadrat Test von Häufigkeiten, Kruskal-Wallis-Test zur Rangdatenanalyse) ergänzt wurden. Die Analyse der Daten mit einem binär-logistischen Regressionsverfahren brachte dabei mit sich, dass die zugrundeliegenden Daten im Rahmen der Codierung für die Analyse stark vereinfacht werden mussten, da nur dichotome Merkmalsausprägungen im Rahmen einer derartigen Analyse berücksichtigt werden konnten. Um sämtliche für die Beantwortung der zugrundeliegenden Fragestellung notwendigen Variablen allerdings in einer Einzelanalyse zusammenfassen zu können und die Problematik multipler Einzeltestungen damit zu umgehen wurde dieses Vorgehen trotz der soeben beschriebenen Einschränkung als sinnvoll erachtet. Zur Vorbereitung der Regressionsanalyse wurde ausschließlich die Vergleichsgruppe auf dem am meisten präferierten (*ersten*) Rangplatz ohne Berücksichtigung des Targetgeschlechts nach der Richtung des Vergleichs codiert (0 = Aufwärtsvergleich, Uni oder Gymnasium; 1 = Abwärtsvergleich, Real- oder Hauptschule) und als abhängige Variable in das Regressionsmodell eingefügt. Als Prädiktoren flossen die Faktoren Untersuchungsbedingung (0 = Kontrollgruppe, 1 = Experimentalgruppe), das Geschlecht der Versuchsperson (0 = weiblich, 1 = männlich) und der Interaktionsterm Untersuchungsbedingung \times Geschlecht in die Regressionsanalyse ein. Die Ergebnisse der

Analyse zeigen mit $R^2 = .11$ einen signifikanten Einfluss der Untersuchungsbedingung, $B = -2.49$, $SE(B) = 1.14$, $Wald = 4.78$, $p < .05$ in Richtung einer geringeren Tendenz zu Abwärtsvergleichen in der Experimentalgruppe. Ein tendenziell signifikanter Einfluss des Prädiktors Geschlecht macht deutlich, dass Jungen eher eine geringere Tendenz zu Abwärtsvergleichen aufweisen als Mädchen, $B = -1.45$, $SE(B) = .899$, $Wald = 2.59$, $p = .11$. Ergänzt werden die Ergebnisse durch einen signifikanten Interaktionseffekt, aus welchem hervorgeht, dass Jungen der Experimentalbedingung eine im Vergleich zu den anderen Gruppen erhöhte Tendenz zu Abwärtsvergleichen zeigen, $B = 3.75$, $SE(B) = 1.46$, $Wald = 6.59$, $p = .01$. Ergebnisse einer post-hoc Analyse mit dem Chi-Quadrat Test getrennt nach Geschlecht ergänzten die regressionsanalytisch nachgewiesenen Befunde, wenngleich festzuhalten blieb, dass der Großteil der befragten Probanden/-innen (83.72% aller Mädchen und 80.56% aller Jungen) einen sozialen Vergleich nach höchster leistungsmäßiger Ähnlichkeit (Leistungsmaßstab Gymnasium oder höher) anstellen wollte, wobei der Anteil von Jungen ($n = 12$; 41.4%) in der Experimentalgruppe geringer ausfällt, als in der Kontrollgruppe ($n = 17$; 58.6%) und sich bei Mädchen ein entgegengesetztes Befundmuster zeigt ($n_{EG} = 24$; 66.7% / $n_{KG} = 12$; 33.3%), $\chi^2(1, N = 65) = 4.16$, $p < .05$.

Die Ergebnisse der weiterführenden Analysen sind in Tabelle 11 zusammengefasst. Es wird ersichtlich, dass entgegen der Ergebnisse der Regressionsanalyse hier zwischen Jungen der Experimental- und Kontrollgruppe nur ein schwach tendenziell signifikanter Unterschied in Richtung einer höheren Rate von Abwärtsvergleichen in der Experimentalgruppe vorlag, $\chi^2(1, N = 36) = 2.04$ $p = .15$, wohingegen bei Mädchen ein signifikanter Effekt verdeutlicht, dass Mädchen der Experimentalgruppe deutlich weniger häufig einen Abwärtsvergleich präferierten, $\chi^2(1, N = 43) = 6.61$ $p = .01$. Es scheint demnach so zu sein, dass der signifikante Interaktionsterm aus der Regressionsanalyse insbesondere durch die geringe Anzahl von Abwärtsvergleichen bei Mädchen der Experimentalbedingung gespeist wurde. Zu berücksichtigen bleibt bei der Chi-Quadrat-Analyse, dass durch die geringe Anzahl von Abwärtsvergleichen und einer damit einhergehenden Verminderung der erwarteten Häufigkeiten unter den geforderten, kritischen Wert von 5 die Ergebnisse nur mit Vorsicht interpretiert werden dürfen.

Tabelle 11: Chi-Quadrat-Analysen und Häufigkeitsverteilung der sozialen Vergleiche (dichotom: aufwärts/parallel vs. abwärts) von Jungen und Mädchen in Abhängigkeit der Untersuchungsbedingung.

	Jungen		Mädchen	
	Parallel- & Aufwärtsvergl.	Abwärtsvergleich	Parallel- & Aufwärtsvergl.	Abwärtsvergleich
Experimentalgruppe	12	5	24	1
Kontrollgruppe	17	2	12	6
	$\chi^2(1, N = 36) = 2.05, p = .15$		$\chi^2(1, N = 43) = 6.61, p = .01$	
Gesamt:	$\chi^2(1, N = 79) = .73, p = .39$			

Wird die Kategorisierung der Vergleichsdimension nicht dichotom, sondern nach Aufwärts- (Uni), Parallel- (Gymnasium) und Abwärtsvergleich (Realschule und Hauptschule) vorgenommen zeigten sich signifikante Unterschiede in den Häufigkeiten analog zu den o.g. Befunden (Tabelle 12): Mädchen der Experimentalgruppe präferierten deutlich häufiger den Vergleich mit einer statushöheren und seltener den Vergleich mit einer statusniedrigeren Gruppe als Mädchen der Kontrollgruppe. Der nur tendenziell signifikante Vergleich zwischen Jungen der Experimental- und Kontrollbedingung aus der vorausgegangenen Analyse lag nur noch in abgeschwächter Form vor.

Tabelle 12: Chi-Quadrat-Analysen und Häufigkeitsverteilung der sozialen Vergleiche (trichotom: aufwärts vs. parallel vs. abwärts) von Jungen und Mädchen in Abhängigkeit der Untersuchungsbedingung.

	Jungen			Mädchen		
	Aufwärtsvergleich	Parallelvergleich	Abwärtsvergleich	Aufwärtsvergleich	Parallelvergleich	Abwärtsvergleich
Experimentalgruppe	6	6	5	10	14	1
Kontrollgruppe	5	12	2	4	8	6
	$\chi^2(2, N = 36) = 3.28, p = .19$			$\chi^2(2, N = 43) = 6.82, p < .05$		
Gesamt:	$\chi^2(2, N = 79) = 1.94, p = .38$					

Werden alle zur Verfügung stehenden Vergleichsgruppen (Studienanfänger/-innen, Gymnasium, Realschule, Hauptschule) einer Häufigkeitsanalyse unterzogen marginalisierten sich die bisher bestehenden Effekte zunehmend i.S.v. einer Tendenz geringer Häufigkeiten von Abwärtsvergleichen und einer häufigeren Präferenz von Aufwärtsvergleichen bei Mädchen der Experimentalgruppe im Vergleich zu Mädchen der Kontrollgruppe. Bei Jungen divergierten die Häufigkeiten hinsichtlich der Präferenz für die verschiedenen sozialen Vergleichsgruppen nicht in Abhängigkeit der Untersuchungsbedingung (Tabelle 13).

Tabelle 13: Chi-Quadrat-Analysen und Häufigkeitsverteilung der sozialen Vergleiche (vierfach gestuft nach Schulform) von Jungen und Mädchen in Abhängigkeit der Untersuchungsbedingung.

	Jungen			
	Studierende	Gymnasium	Realschule	Hauptschule
Experimentalgruppe	6	6	1	4
Kontrollgruppe	5	12	0	2
$\chi^2 (3, N = 36) = 3.66, p = .30$				
	Mädchen			
	Studierende	Gymnasium	Realschule	Hauptschule
Experimentalgruppe	10	14	0	1
Kontrollgruppe	4	8	1	5
$\chi^2 (3, N = 43) = 6.92, p = .07$				
Gesamt:	$\chi^2 (3, N = 79) = 1.99, p = .58$			

Ferner wird ersichtlich, dass im Rahmen von Abwärtsvergleichen nur ein Mädchen (5.6 %) der Kontrollgruppe und ein Junge (5.9%) der Experimentalgruppe einen Vergleich mit Realschüler/-innen gewünscht haben. Der Großteil der verbleibenden Abwärtsvergleiche (1 Mädchen und 4 Jungen der Experimentalgruppe sowie 5 Mädchen und 2 Jungen der Kontrollgruppe) ging auf die Vergleichsgruppe der Hauptschüler/-innen zurück und indiziert somit, dass bei den vorausgegangenen Analysen (in welchen Real- und Hauptschule in eine Vergleichsgruppe zusammengefasst wurden) der statusniedrigsten Vergleichsgruppe die größte Bedeutung im Rahmen der abwärtsgerichteten Vergleiche zugekommen zu sein scheint.

Bisher wurden ausschließlich die am meisten präferierten Vergleichsgruppen (Rangplatz 1) ausgewertet. Eine differenziertere Betrachtung über alle abgefragten Rangplätze und Vergleichsgruppen hinweg erlaubt die aufbauende Analyse mit dem Kruskal-Wallis-Test für Rangdaten. Die Ergebnisse der Rangdaten ohne Einbezug des Targetgeschlechts verdeutlichten in ersten Analysen, dass zwischen den Untersuchungsbedingungen unabhängig vom Geschlecht der Versuchsperson keine Unterschiede hinsichtlich der Rangplätze der vier verschiedenen sozialen Vergleichsgruppen festzustellen war, alle $H < 1.80, p > .18$. Analysen getrennt nach Geschlecht wiesen darauf hin, dass Jungen ($RM = 3.02$) sich tendenziell weniger gern als Mädchen ($RM = 2.76$) mit der Gruppe von Realschüler/-innen verglichen, $H = 3.15, p = .08$. Bei allen anderen sozialen Vergleichsgruppen divergierten die restlichen drei Rangplätze nicht in Abhängigkeit des Geschlechts der Versuchsperson, alle $H < 2.15, p > .15$. Von besonderer Bedeutung für die Beantwortung der dieser Studie zugrundeliegenden Fragestellung waren jedoch die Einzelvergleiche der beiden Geschlechtergruppen in Abhängigkeit von der Untersuchungsbedingung. Für die Gruppe der Jungen konnten keine systematischen Unterschiede auf den Rangplätzen für alle vier angebotenen Vergleichsgruppen gefunden werden, alle $H < 1.71, p > .19$. In der Gruppe der Mädchen und in Bezug auf die Vergleichsgruppe der Hauptschüler/-innen fand sich jedoch

ein marginal signifikanter Unterschied in Richtung einer höheren Rangplatzzuweisung dieser Vergleichsgruppe von Mädchen der Experimentalgruppe ($RM = 3.49$) im Vergleich zu einem geringeren Rangplatz bei Mädchen in der Kontrollbedingung ($RM = 2.78$). Mädchen der Experimentalbedingung scheinen somit, nachdem Sie das Material verarbeitet haben, aus welchem die schlechtere Stellung von Jungen im Bildungssystem hervorgeht, weniger gern den Vergleich mit der statusniedrigsten sozialen Gruppe anstellen zu wollen, $H = 3.54$, $p = .06$. Bei den drei verbleibenden Vergleichsgruppen fanden sich keine weiteren, bedeutsamen Unterschiede zwischen Mädchen der Experimental- vs. Kontrollgruppe, alle $H < 1.53$, $p > .22$. Die mittleren Rangplätze der angebotenen sozialen Vergleichsgruppen in Abhängigkeit von Geschlecht und Untersuchungsbedingung fasst Abbildung 14 grafisch zusammen.

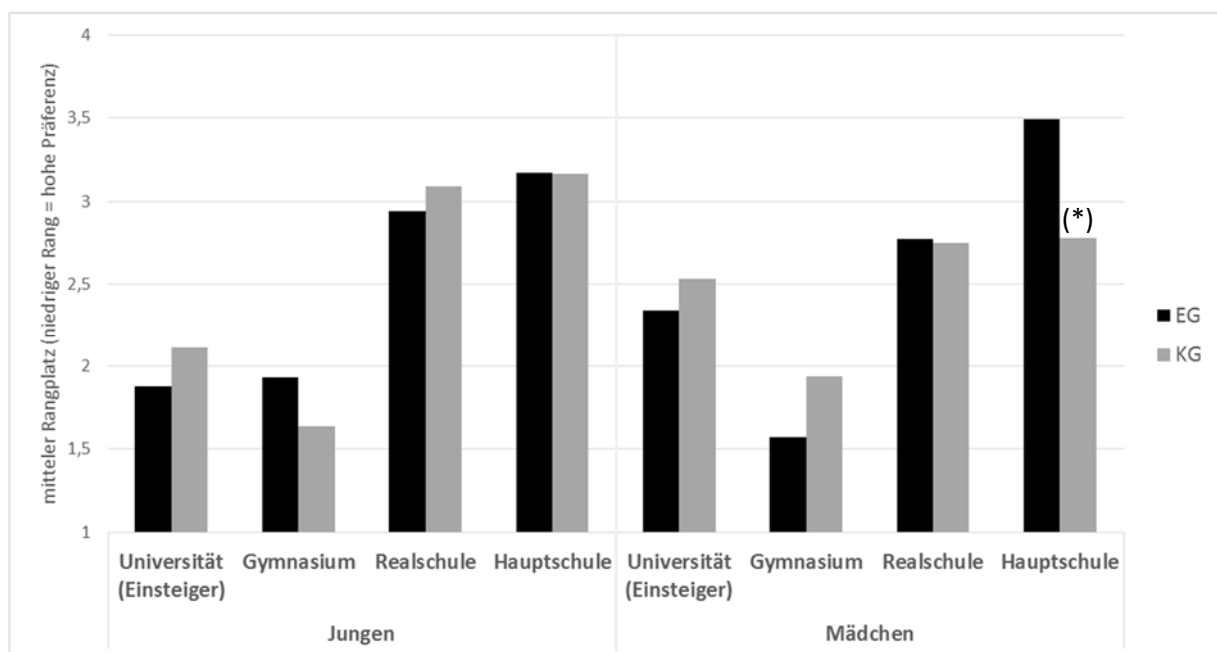


Abbildung 14: Mittlere Rangplätze für den sozialen Vergleich (vierfach gestuft, nach Schulform) getrennt nach Geschlecht und Untersuchungsbedingung. (*) = $p < .10$.

Wird in die Analyse der Rangdaten abschließend auch das Targetgeschlecht einbezogen, d.h. berücksichtigt, mit welchem Geschlecht innerhalb einer Vergleichsgruppe die eigene Leistung am ehesten kontrastiert werden soll, zeigte sich, dass unabhängig von der Untersuchungsbedingung das eigene Geschlecht häufig am meisten präferiert wurde. So konnte festgestellt werden, dass Jungen sich beim Aufwärtsvergleich (Studienanfänger/-innen) deutlich lieber mit einer männlichen Zielperson verglichen ($RM = 2.92$) als Mädchen ($RM = 4.72$), $H = 12.70$, $p < .001$. Gleiches galt für den Parallelvergleich (Gymnasium: $RM_{\text{Jungen}} = 2.56$; $RM_{\text{Mädchen}} = 3.63$; $H = 7.13$, $p < .01$) und tendenziell auch für den stärksten Abwärtsvergleich (Hauptschule: $RM_{\text{Jungen}} = 5.63$, $RM_{\text{Mädchen}} = 6.26$; $H = 2.43$, $p = .12$). Bei Mädchen fand sich ein ähnlich gelagertes Bild: Während beim Aufwärtsvergleich keinerlei Präferenz zugunsten der eigenen Geschlechtergruppe festzustellen war, wurde beim Parallelvergleich der Vergleich mit einer weiblichen Zielperson

von Mädchen deutlich stärker bevorzugt ($RM = 2.29$) als von Jungen ($RM = 3.42$), $H = 9.63$, $p < .01$. Auch beim abwärtsgerichteten Vergleich mit Realschüler/-innen wünschten Mädchen ebenso lieber den Vergleich mit einer weiblichen Targetperson als mit Jungen ($RM_{\text{Jungen}} = 5.92$; $RM_{\text{Mädchen}} = 4.51$; $H = 15.88$, $p < .001$), wie tendenziell beim Abwärtsvergleich mit der statusniedrigsten Vergleichsgruppe (Hauptschule: $RM_{\text{Jungen}} = 6.03$; $RM_{\text{Mädchen}} = 5.52$; $H = 2.55$, $p = .11$). Eine Präferenz für den Vergleich mit der eigenen Geschlechtergruppe fand sich somit in allen abgefragten Vergleichsmöglichkeiten, mit Ausnahme von Mädchen beim Aufwärtsvergleich und Jungen beim abwärtsgerichteten Vergleich mit Realschülern.

Analysen getrennt nach Geschlecht zeigten, dass in Abhängigkeit von der Untersuchungsbedingung bei Jungen lediglich ein marginal bedeutsamer Unterschied festgestellt werden konnte. Jungen der Experimentalgruppe vergaben für die Vergleichsgruppe der männlichen Studienanfänger einen tendenziell geringeren Rangplatz ($RM = 2.50$) als Jungen der Kontrollbedingung ($RM = 3.32$), $H = 3.24$, $p = .07$. Hinsichtlich aller anderen Vergleichsdimensionen ergab diese Analyse keine weiteren bedeutsamen Unterschiede, alle $H < 1.09$, $p > .30$. Bei Mädchen zeigten sich tendenziell signifikante Differenzen bei der Vergleichsgruppe weiblicher Studienanfängerinnen in Richtung einer höheren Präferenz dieser Vergleichsgruppe in der Experimentalbedingung ($RM_{\text{EG}} = 3.16$, $RM_{\text{KG}} = 4.28$, $H = 2.38$, $p = .12$) sowie in gleicher Stärke und Richtung bei der Vergleichsgruppe der Gymnasiastinnen ($RM_{\text{EG}} = 1.92$, $RM_{\text{KG}} = 2.78$, $H = 2.21$, $p = .14$). Lediglich bei der Vergleichsgruppe der Hauptschülerinnen wurde der vergleichsweise höhere Rangplatz (entspricht einer geringeren Präferenz) bei Mädchen der Experimentalgruppe ($RM = 6.96$) im Vergleich zu Mädchen der Kontrollgruppe ($RM = 5.28$) statistisch bedeutsam, $H = 6.60$, $p = .01$. Alle weiteren abgefragten Vergleichsgruppen unterschieden sich innerhalb der Gruppe der Mädchen nicht in Abhängigkeit von der Untersuchungsbedingung, alle $H > 1.50$, $p < .21$.

11.4. Diskussion

Die Ergebnisse der vorliegenden Studie weisen ein sehr heterogenes Befundmuster auf. In der initialen logistischen Regressionsanalyse kann zwar ein Effekt dahingehend festgestellt werden, dass Jungen in der Experimentalgruppe deutlich häufiger zu Abwärtsvergleichen auf dem ersten Rangplatz tendieren als die anderen Vergleichsgruppen, wobei festgehalten werden muss, dass das Niveau der vorliegenden Daten zur Erfüllung der Voraussetzungen eines binär-logistischen Regressionsverfahrens stark vereinfacht (dichotomisiert) wurde. Vertiefende Analysen mit verteilungsfreien Verfahren deuten hingegen darauf hin, dass a) sowohl Jungen als auch Mädchen am ehesten einen Vergleich mit der eigenen Schulform oder höher anstellen möchten sowie b) insbesondere in der Gruppe der Mädchen ein bedeutsamer Unterschied in der Tendenz zu einer geringeren Wahrscheinlichkeit von Abwärtsvergleichen auf dem ersten Rang-

platz bei Mädchen der Experimentalbedingung im Vergleich zur Kontrollbedingung vorgefunden werden konnte. Unter Einbezug aller zu vergebenden Rangplätze weisen die Ergebnisse aus den Analysen auf nur sehr wenige und meist nur schwach signifikante Unterschiede hin, so z.B. dass Mädchen der Experimentalgruppe sich tendenziell weniger gern mit der Gruppe der Hauptschüler/-innen vergleichen möchten, als Mädchen der Kontrollgruppe. Bei Jungen hingegen lassen sich wiederum keine signifikanten Unterschiede in den Rangplätzen für den sozialen Vergleich in Abhängigkeit der Untersuchungsbedingung aufzeigen. Unter Einbezug des Targetgeschlechts, also der Geschlechtergruppe innerhalb der sozialen Vergleiche, konnte eine unabhängig von der Experimentalbedingung stattfindende Präferenz für die eigene Geschlechtergruppe in nahezu allen angebotenen sozialen Vergleichsgruppen (zumindest tendenziell) festgestellt werden. Trennt man diese Analysen nach Geschlecht wird ersichtlich, dass Jungen tendenziell eher einen Aufwärtsvergleich (mit Jungen zu Beginn des Studiums) anstreben. Bei Mädchen zeigt sich ergänzend, dass die Experimental- im Vergleich zur Kontrollgruppe eher Aufwärts- und Parallelvergleiche und seltener Abwärtsvergleiche mit Mädchen der entsprechenden Vergleichsgruppen anstreben. Es kann somit festgehalten werden, dass die Befunde der Regressionsanalyse eher auf die Unterschiede von Mädchen zwischen Experimental- vs. Kontrollgruppe zurückzuführen sind, da sich bei Jungen kaum bedeutsame Unterschiede zwischen den Untersuchungsbedingungen in den Einzelvergleichen (Häufigkeits- und Rangdatentests) zeigen.

Für die aufgestellten und zu überprüfenden Hypothesen bedeuten diese Ergebnisse, dass Hypothese 18 nicht beibehalten werden kann. Zwar deuten die Ergebnisse der Regressionsanalyse darauf hin, dass Jungen der Experimentalbedingung verstärkt auf Abwärtsvergleiche zurückgreifen, allerdings wird dieser Befund durch zahlreiche Gegenevidenz in den vertiefenden Analysen mit nichtparametrischen Verfahren konterkariert. Vielmehr scheint es so, dass durch die Bearbeitung des experimentellen Treatments verstärkt Mädchen (und nicht Jungen) in der Wahl ihrer am meisten präferierten Vergleichsgruppe beeinflusst werden und nach der Bearbeitung des Materials über schulversagende Jungen weniger bereit sind, abwärtsgerichtete Vergleiche anzustellen. Die Gruppe der Jungen hingegen bleibt von der experimentellen Bedrohung der Eigengruppe scheinbar weitgehend unbeeinflusst. Allein zu Hypothese 19 findet sich umfangreichere Evidenz dahingehend, dass in der Tat wie erwartet von Jungen wie auch von Mädchen in beiden Untersuchungsbedingungen die eigene Geschlechtergruppe bevorzugt für den sozialen Vergleich herangezogen wird und damit die generelle Wichtigkeit des eigenen Geschlechts als zentrale Kategorisierungsdimension für soziale Vergleichsprozesse im Jugendalter unterstreicht. Zusätzlich konnte gezeigt werden, dass zwar die größte Gruppe von Mädchen (51.2%) unabhängig von der Experimentalbedingung den Vergleich mit der eigenen Schulform präferiert, jedoch ist dasselbe Befundmuster bei Jungen vorzufinden, welche mit 50.0% ebenfalls zum Großteil die eigene Schulform für den sozialen Vergleich bevorzugen. Zusätzlich divergieren die Präferenzen

für den sozialen Vergleich bei Mädchen hinsichtlich Auf- und Abwärtsvergleiche (mehr Aufwärtsvergleiche und weniger Abwärtsvergleiche bei Mädchen der Experimentalgruppe) in Abhängigkeit des experimentellen Treatments. Somit kann insbesondere im Hinblick auf den Vergleich mit der Gruppe der Jungen nicht davon ausgegangen werden, dass sich hauptsächlich Mädchen eher mit der eigenen Schulform vergleichen wollen.

Die vorliegenden Befunde stehen im Kontrast zu einer Reihe empirischer Untersuchungen, welche den Einsatz und die Wirksamkeit von Abwärtsvergleichen als Reaktion auf eine Bedrohung der sozialen Identität nachweisen (Buunk, Collins, Taylor, Van Yperen & Dakof, 1990; Gibbons & Gerrard, 1989; Suls, Martin & Wheeler, 2002; Wheeler & Myake, 1992; Wood, Taylor & Lichtman, 1985) und ferner nachweisen, dass ein Bedürfnis nach Erhöhung des eigenen Selbstwerts in Reaktion auf eine Bedrohungssituation abwärtsgerichtete Vergleiche motiviert, wohingegen ein Bedürfnis nach Bewertung der eigenen Person und eine Verortung im sozialen Gefüge maßgeblich zu einem Einsatz von Aufwärtsvergleichen beiträgt (Collins, 1996; Wills, 1991). Taylor & Lobel (1989) ergänzen diese Befundlage und subsumieren, dass zwei Mechanismen sozialen Vergleichsprozessen zugrunde liegen und infolge einer Bedrohung des eigenen Selbst relevant werden: Einerseits stellen abwärtsgerichtete Vergleiche den bedrohten Selbstwert wieder her bzw. unterstützen diesen, durch die Verletzung der Fairness dieses Vergleichs sorgen hingegen aufwärtsgerichtete Explorationen („*upward contacts*“) für eine Steigerung der Motivation zur Änderung des relativ schlechteren Vergleichsergebnisses und erhöhen die Hoffnung auf eine Verbesserung des Vergleichsergebnisses. Verschiedene Arbeiten tragen zusätzlich zur Erklärung bei, warum in der vorliegenden Untersuchung Jungen nach Bedrohung ihrer sozialen Identität nicht geneigt sind, durch Abwärtsvergleiche die positive Distinktheit des Selbst und der Eigengruppe wiederherzustellen. So existieren Hinweise darauf, dass Gruppen, die als unterlegen wahrgenommen oder ebenso dargestellt werden, eher Intragruppenvergleiche präferieren, als Intergruppenvergleiche anzustellen (Lockwood, 2002; Zagefka & Brown, 2005). Hierfür sprechen die vorliegenden Ergebnisse zum Teil, da Jungen eher den Vergleich mit der Eigengruppe suchen, wenngleich zu berücksichtigen bleibt, dass dieses Befundmuster unverändert auch bei Mädchen vorzufinden ist, welche zu keinem Zeitpunkt der Untersuchung einer Bedrohungssituation ausgesetzt waren. Weiterhin bleibt zu diskutieren, dass die Richtung des Vergleichs (auf- vs. abwärts) nicht automatisch und deterministisch mit einer eindeutig bestimmten (affektiven) Konsequenz verbunden ist, d.h. nicht jeder Abwärtsvergleich zieht einen als positiv wahrgenommenen *Outcome* nach sich (Buunk et al., 1990; Smith, 2000). Dies kann erklären, warum ein Abwärtsvergleich auch als bedrohlich gewertet und in der Konsequenz eher vermieden werden kann, da ein abwärtsgerichteter Vergleich keinerlei Garantie für eine emotional positive Konsequenz bietet. Dies hängt vielmehr von zahlreichen korrespondierenden Faktoren ab, wie z.B. Selbstbild und Selbstwert, wahrgenommene Ähnlichkeit mit der Vergleichsgruppe

und Intergruppenkontakt mit der Vergleichsgruppe (Zagefka & Brown, 2005). Es kann somit diskutiert werden, ob der Vergleich mit einer anderen, niedrigeren und auch der niedrigsten Schulform bei den untersuchten Jungen, welche der höchsten deutschen Schulform angehören, zu bedrohlich (im Sinne von zu unähnlich) wahrgenommen wird, als dass bei einem Vergleich mit diesen entsprechenden Gruppen positive Konsequenzen durch den Abwärtsvergleich erwartet werden. Weiterhin weisen Taylor et al. (1996) darauf hin, dass eine langfristige Exposition einer Bedrohung die Wirksamkeit von Abwärtsvergleichen einschränkt und diese emotional eher negativ besetzt erscheinen lassen, wodurch diese weniger attraktiv erscheinen und ein Einsatz als Strategie zur Wiederherstellung positiver Distinktheit unwahrscheinlicher wird. Die in der vorliegenden Studie eingesetzte experimentell induzierte Bedrohung wurde zwar kurzfristig angelegt und hatte somit ein hohes Potential für den Einsatz und die Wirksamkeit eines Abwärtsvergleichs, jedoch kann nicht zweifelsfrei festgestellt werden, inwiefern eine ähnliche Bedrohung im Schulalltag häufiger anzutreffen ist (vgl. *Studie 1* und Kapitel 1). Insofern kann nicht ausgeschlossen werden, dass u.U. doch von einer längerfristigen Bedrohung auszugehen ist (oder evtl. auch die Kontrollgruppe einem gewissen Maß an Bedrohung der sozialen Identität unterliegt) und damit einhergehend abwärtsgerichtete Vergleiche u.U. nicht mehr als adäquate Reaktion der Probanden/-innen auf die Bedrohung der sozialen Identität angesehen (vgl. Abs. 7.1.) bzw. auch die Vergleichswahlen der Kontrollgruppe beeinflusst werden. Auch bleibt darauf hinzuweisen, dass der von Jungen präferierte Vergleich mit anderen Jungen derselben Schulform eigentlich ebenso als eine leichte Form eines Abwärtsvergleichs zu klassifizieren ist, da Jungen sich somit auch mit einer eher unterlegenen Gruppe vergleichen. Inwiefern dies allerdings durch das Bedürfnis nach maximaler Ähnlichkeit sozialer Vergleichsgruppen konterkariert wird, kann im Rahmen der vorliegenden Studie nicht beantwortet werden.

Zu erklären bleibt abschließend, warum aufgefundene Befunde ausschließlich auf die Gruppe der Mädchen zurückzuführen waren. Einerseits scheint bedeutsam, dass insbesondere – auch im Hinblick auf längsschnittliche Verfestigungen sozialer Vergleichsmuster – der sozialen Vergleichsorientierung eine bedeutende Rolle zukommt (Buunk, Zurriaga, Gonzalez-Roma & Subirats, 2003). Bei weiblichen Probanden wurde eine geringe, jedoch signifikant höhere Ausprägung dieser Orientierung nachgewiesen (Gibbons & Buunk, 1999). Für das Jugendalter liegen keine Vergleichswerte vor, allerdings kann bezüglich der vorliegenden Ergebnisse angenommen werden, dass in dieser Altersspanne möglicherweise noch größere Geschlechterunterschiede vorherrschend sind. Hinzu kommt, dass Jungen sich als sozial unabhängiger wahrnehmen und beschreiben, wohingegen das Selbst von Mädchen stärker sozial dependent ausgeprägt ist (Guimond, Chatard, Martinot, Crisp & Redersdorff, 2006). Es scheint denkbar, dass diese Art der Wiederherstellung positiver Distinktheit eher für Mädchen geeignet zu sein scheint, auch im Hinblick auf deren höhere Motivation in der Schule (Martin, 2004; Stamm, 2007). Da durch die

Aufwärtsvergleiche als Reaktion auf die Darstellung der Gruppe der Jungen als schulisch unterlegen – welche Mädchen der Experimentalgruppe deutlich häufiger anstellen wollten, als Mädchen der Kontrollgruppe – neben einer Erhöhung des Selbst(werts) auch eine Erhöhung der Motivation verbunden mit dem Streben nach Selbstverbesserung einhergeht (Blanton, Buunk, Gibbons & Kuyper, 1999; Collins, 1996; Huguet, Dumas, Monteil & Genestoux, 2001), könnte das durch Mädchen (der Experimentalgruppe) gezeigte Befundmuster dazu beitragen, die für Mädchen positiven Unterschiede und damit den Überlegenheitsstatus aufrecht zu erhalten.

Limitationen. Wie in den vorausgegangenen Studien muss kritisch diskutiert werden, inwiefern die ausbleibenden treatmentbezogenen Effekte in der vorliegenden Studie durch eine ungünstige Operationalisierung der abhängigen Variable (Vergleichsgruppen) mitverursacht werden. Wie dargestellt wurde hängt die Instrumentalität von Abwärtsvergleichen maßgeblich von der Art der und der Beziehung sowie dem Kontakt zur Vergleichsgruppe ab (Zagefka & Brown, 2005). Auf Grundlage der gefundenen Ergebnisse und unter der Annahme, dass zwischen verschiedenen Schulformen kaum Intergruppenkontakte bestehen kann gemutmaßt werden, dass ein schulartübergreifender Vergleich wie in der vorliegenden Studie u.U. nicht angemessen ist, da ein Abwärtsvergleich mit einer niedrigeren Schulform eine nur schlecht applikable und ggf. zu bedrohliche Reaktion auf den ursprünglichen Bedrohungsreiz darstellt. Damit einhergehend kann kein positiver *Outcome* eines derart ausgerichteten Vergleichs antizipiert werden. Zur Vermeidung dieser Problematik wäre denkbar, in weiteren diesbezüglichen Erhebungsettings auf einen Vergleich innerhalb der Klasse (mit sehr gut, durchschnittlich und schlecht leistenden Schüler/-innen) oder der Schule (mit einer gut, neutral oder schlecht leistenden Klasse dieser Schule) zurückzugreifen, um die Bedrohlichkeit eines Abwärtsvergleichs abzuschwächen. Weiterhin fand die Erhebung an der höchsten deutschen Schulform (Gymnasium) statt, auch um ein möglichst breites Spektrum an Abwärtsvergleichen zu ermöglichen. Inwiefern diese Tatsache allerdings schwächend auf das experimentelle Treatment wirkt oder gar demgegenüber insensitiv macht, da in dieser Schulform der höchstmögliche Abschluss angestrebt wird und somit die Probanden/-innen einem sehr bildungserfolgreichen Klientel entsprechen, kann im Rahmen der vorliegenden Studie nicht genau beantwortet werden. Ferner ist nicht eindeutig ersichtlich, inwiefern durch den Schulalltag und die Einflüsse von signifikanten Dritten sowie der medialen Berichterstattung über Jungen davon ausgegangen werden muss, dass auch Jungen der Kontrollbedingung einem gewissen Niveau an Bedrohung der sozialen Identität unterliegen. In zukünftigen Studien wäre es hilfreich, eine Ausgangslagemessung vorzunehmen, um genauer feststellen zu können welches grundlegende Ausmaß an sozialer Bedrohung vorhanden ist und wie dieses sich durch die Bearbeitung des Treatment- vs. Kontrollmaterials verstärkt oder nicht.

Zusätzlich wirkte sich das aufgrund der zeitlichen und administrativen Beschränkungen vergleichsweise einfache Design der Studie mindernd auf die Aussagekraft der Ergebnisse aus, insbesondere im Hinblick auf die Erfassung weiterer Variablen, welche Richtung und Stärke möglicher Effekte beeinflussen können, so z.B. soziale Vergleichsorientierung, schulisches Selbstkonzept oder Wert von Schule und Bildung. Ein verstärkter Einbezug solcher Variablen würde es ermöglichen, die angestellten Analysen genauer zu spezifizieren und mögliche Moderations- oder Mediationseffekte aufzudecken.

Abschließend wurde die Vergleichswahl nur generell abgefragt, wohingegen Aufgaben in zwei Schulfächern (Deutsch, Mathematik) zu bearbeiten waren. Um differenzielle Effekte besser abbilden und der Tatsache Rechnung zu tragen, dass Jungen in einigen Schulfächern Leistungsvorteile aufweisen wäre es hilfreich, die Wahl der Vergleichsgruppe zukünftig fachspezifisch zu erheben, um somit feststellen zu können, ob Jungen in männlich konnotierten Leistungsdomänen zwar keinen Bedarf für Abwärtsvergleiche sehen, jedoch in typisch weiblich konnotierten Fachbereichen abwärtsgerichtete Vergleiche als funktional für die Wiederherstellung positiver Distinktheit in Abhängigkeit von der Experimentalbedingung ansehen.

III. GESAMTDISKUSSION

Ziel der vorliegenden Arbeit war es zu identifizieren, ob und in welchem Ausmaß Jungen durch die aufkommende mediale Debatte über ihr angebliches Scheitern in der Schule und die daraus erwachsenden negativen Konsequenzen in ihrer Leistungsfähigkeit und schulischen Motivation im Sinne eines *stereotype threat* (Steele & Aronson, 1995) beeinträchtigt werden. Ferner war von Interesse, welche Strategien Jungen in Anlehnung an die Theorie der sozialen Identität (Tajfel & Turner, 1986) einsetzen um (emotional) negative Konsequenzen dieses Diskurses zu vermeiden oder auszugleichen.

In einem ersten Schritt blieb anhand statistischer Daten und der Ergebnisse der großen Schulleistungsstudien zu überprüfen, inwiefern die mediale Darstellung von Jungen als Bildungsverlierer empirisch verifiziert werden kann. Eine initiale Analyse der vorhandenen wissenschaftlichen und medialen Aufarbeitung der Jungenproblematik machte deutlich, dass das medial transportierte Bild über Jungen als *generelle* Versager im Bildungssystem nicht vollumfänglich abgesichert werden konnte. Einerseits deuten die Verteilungen von Mädchen und Jungen auf die verschiedenen Schulformen darauf hin, dass Jungen in niedrigen Schulformen (z.B. Haupt- und Förderschulen) stark überrepräsentiert sind, wobei Mädchen auf der höchsten deutschen Schulform (Gymnasium) den größeren Anteil stellen, damit einhergehend Jungen ebenfalls eher geringer wertige schulische Abschlusszertifikate erwerben und die Schule häufiger ohne Abschluss verlassen (Statistisches Bundesamt, 2014c, 2014d). Andererseits deutet ein Blick auf den Kompetenzerwerb an, dass Jungen keineswegs durchweg schlechtere Kompetenzen erzielen. Der Rückstand in der Lesekompetenz auf Mädchen fällt je nach Anlage der Leistungsmessung und der untersuchten Altersgruppe verschieden groß, jedoch stets signifikant zugunsten der Mädchen aus (vgl. auch Hannover & Kessels, 2011). Bereits im Grundschulalter kann ein Leistungsvorsprung von Mädchen festgestellt werden (Bos, et al., 2012a), welcher sich im Jugendalter sehr stark zu einem Kompetenzunterschied von ca. 2 Schuljahren zugunsten der Leseleistung von Mädchen ausweitet (Prenzel et al., 2013). Bei den mathematischen Kompetenzen hingegen ergibt sich bereits im Grundschulalter ein signifikanter Leistungsvorsprung von Jungen (Bos et al., 2012b), welcher auch bei 15jährigen Jugendlichen in der PISA-Studie in gleichem Maße aufrechterhalten wird (Prenzel et al., 2013). Leistungsvorteile von Jungen im Grundschulalter im naturwissenschaftlichen Bereich lassen sich hingegen im Jugendalter nicht mehr feststellen. Zusätzlich weisen Jungen in der PISA-Untersuchung 2012 eine höhere Problemlösekompetenz auf als Mädchen.

Insgesamt bleibt zwar zu berücksichtigen, dass Jungen durch ihre vergleichsweise geringere Repräsentanz in höheren Schulformen auch Nachteile im Kompetenzerwerb erfahren, jedoch kann nicht von einem generellen Versagen von Jungen in der Schule ausgegangen wer-

den. Vielmehr zeichnet sich empirisch ein Bild was darauf hindeutet, dass die Geschlechterunterschiede sehr fächerspezifisch und entsprechend der gängigen Geschlechterstereotype über die jeweils untersuchte Fachdomäne ausfallen (Nosek et al., 2002; Steffens & Jelenec, 2011; vgl. auch Kap. 1). Zusätzlich bleibt anzumerken, dass nahezu alle berichteten Kompetenzunterschiede wesentlich geringer ausfallen oder gar verschwinden, wenn um die besuchte Schulform kontrolliert wird (Stanat & Kunter, 2003). Vor dem Hintergrund dieser Diskrepanz zwischen der öffentlichen Berichterstattung über Jungen in der Schule und den empirischen Ergebnissen über den Bildungserfolg von Jungen stellt sich zunächst die initiale Frage, ob Jungen das medial skizzierte Bild über die Eigengruppe antizipieren oder aber wahrnehmen, dass sie sehr wohl in einigen Fachdomänen bessere Leistungen erzielen als Mädchen und sich damit einhergehend nicht als *generelle* Bildungsverlierer betrachten.

Wie Studie 1 dieser Arbeit zeigen konnte, antizipieren die befragten Schülerinnen und Schüler ein gesellschaftlich ungefähr gleich negatives Bild über Jungen und Mädchen im Allgemeinen, wohingegen die Äußerungen mit einem klaren Bezug zum Kontext Lernen und Schule bei Jungen größtenteils negativ ausfallen (z.B. dumm, faul, engagieren sich nicht für Schule, haben schlechte Noten, vergessen Hausaufgaben ständig), während Mädchen im Kontrast fast ausschließlich mit positiven, mit Schulerfolg kompatiblen Eigenschaften und Verhaltensweisen beschrieben werden (z.B. fleißig, strebsam, machen Hausaufgaben, bekommen gute Noten, halten Schulsachen in Ordnung). Auch wenn statt einer offenen Befragung geschlossene Eigenschaftswörterlisten vorgelegt werden bestätigt sich der Befund, dass Jungen im Kontext Schule negativer beschrieben werden, als Mädchen. Ein Effekt des Geschlechts der Versuchsperson in Richtung einer zu erwartenden Eigengruppenbevorzugung (Tajfel, 1982; Mullen et al., 1992) bleibt aus, d.h. sowohl Jungen als auch Mädchen antizipieren dieses Befundmuster gleichermaßen, was ferner für ein Gelingen der Instruktion, es möge eine gesellschaftliche Sichtweise antizipiert und nicht die eigene Meinung dargelegt werden, spricht. Da weiterhin der Anteil von fachspezifischen Aussagen (z.B. begabt für Mathematik) nur einen vernachlässigbar geringen Anteil darstellt, kann subsummiert werden, dass das medial transportierte Stereotyp von *generell* schlechterer Passung von Jungenverhalten zu schulischem Engagement und Erfolg vordergründig antizipiert wird, was so bereits bei einer jüngeren Stichprobe nachzuweisen war (Hartley & Sutton, 2013, *Studie 1*). Diese Ergebnisse erweitern die bisherige Forschungsliteratur um eine globalere Sichtweise. Es wurde nicht überprüft, inwiefern einzelne Fachdomänen eher mit Maskulinität oder Femininität assoziiert werden (z.B. Hyde, Fennema, Ryan, Frost & Hopp, 1990; Nosek et al., 2002; Steffens & Jelenec, 2011), sondern eine generellere Sichtweise auf die Passung zu Schule und schulischen Engagement *per se* eingenommen. Somit können die bisherigen Befunde dahingehend erweitert werden, dass über die gängigen fachbezogenen Geschlechterstereotype hinaus eine schlechtere Passung von Jungen zu Schule und schulischem Engagement generell unabhän-

gig von geschlechtsspezifischen Vor- und Nachteilen in bestimmten Fachdomänen vorzuherrschen scheint. Dies macht deutlich, dass Schüler/-innen trotz der eingangs geschilderten Leistungsvorteile für Jungen in maskulin konnotierten Domänen nicht in der Lage zu sein scheinen, diese auch angemessen in die Wahrnehmung von Jungen im Kontext Schule zu integrieren, sondern eine Überlagerung durch die wahrgenommene schlechtere Passung zur Schule *insgesamt* stattfindet und fachspezifische Vorteile in den Hintergrund drängt. Zusätzlich kann damit deutlich gemacht werden, dass Jungen (und Mädchen) das medial transportierte Bild internalisiert zu haben scheinen. Dies ist insbesondere vor dem Hintergrund relevant, dass bisher unklar war, ob die mediale Berichterstattung, welche sich vor allem auf die Printmedien konzentriert (bei welchen die Nutzung bei Jugendlichen eher rückläufig ist; Klingler, 2008), überhaupt durch die Jugendlichen rezipiert wird, was den Ergebnissen dieser Studie folgend der Fall zu sein scheint. Die Frage nach dem Implementationspfad kann allerdings durch diese Untersuchung nicht beantwortet werden – es bleibt unklar, ob das Erreichen des medialen Diskurses auf direktem oder indirektem Wege (z.B. über Lehrkräfte und deren Erwartungen; Heyder & Kessels, 2012) erfolgt ist.

Aufbauend auf diesem Bild von Jungen und Schule stellt sich die Frage, inwiefern diese Wahrnehmung die Leistungsfähigkeit und Motivation von Jungen beeinflusst. Forschungsarbeiten zur Bedrohung durch Stereotype gehen davon aus, dass ein negatives Stereotyp über die Eigengruppe dazu führen kann, dass durch die Befürchtung dieses Stereotyp durch eigenes Verhalten zu bestätigen kognitive Bewertungs- und Bewältigungsprozesse ausgelöst werden, welche wiederum die Leistungsfähigkeit in der bedrohten Fachdomäne schmälern (*stereotype threat*; Steele & Aronson, 1995; Schmader et al., 2008). Die Auswirkung derartiger Mechanismen auf die Leistung wurden in der bisherigen Forschungsliteratur jedoch zum Großteil auf Mädchen und Frauen in naturwissenschaftlich-technischen Domänen oder im organisationalen Kontext bezogen (u.a. Gupta & Bhawe, 2007; Inzlicht & Ben-Zeev, 2000; Keller, 2002, 2007; für einen Überblick: Nguyen & Ryan, 2008; vgl. auch Kapitel 2). Wenngleich die Ergebnisse einzelner Studien variieren, zeigt sich insgesamt metaanalytisch jedoch eine mittlere Effektstärke von $d = .36$ zuungunsten der Leistung von Frauen unter Stereotypenbedrohung in typischerweise maskulin konnotierten Domänen wie z.B. Mathematik (Nguyen & Ryan, 2008). Nur sehr wenige Studien haben sich nach dem bisherigen Forschungsstand mit der Leistung von Jungen und Männern unter Stereotypenbedrohung in typischerweise feminin konnotierten Domänen auseinandergesetzt (u.a. Hirnstein et al., 2012; Keller, 2007; Koenig & Eagly, 2005). Ein noch geringerer Anteil der Forschung zu *stereotype threat* und Jungen beschäftigte sich mit dem Anwendungskontext Schule und brachte sehr divergierende Befunde hervor. Während Hartley & Sutton (2013) eine generelle Beeinträchtigung in allen untersuchten Leistungsbereichen bei Jungen im Grundschulalter infolge einer Bedrohung durch Stereotype feststellten, fanden Eckert & Imhoff (2013) in

einer Stichprobe 14jähriger Jungen keinerlei leistungsbeeinträchtigende Auswirkungen durch eine Stereotypenbedrohung. Um zu einer weiteren Klärung der Befundlage hinsichtlich der Auswirkung einer Stereotypenbedrohung auf die schulische Leistung von Jungen beizutragen, war es ein weiteres Ziel dieser Arbeit zu überprüfen, inwiefern Jungen durch das in *Studie 1* nachgewiesene Stereotyp im Sinne einer Stereotypenbedrohung in ihrer schulischen Leistungsfähigkeit beeinträchtigt werden. Es konnte gezeigt werden, dass Jungen nach einer Exposition des Stereotyps eines männlichen, generellen Schulversagers ausschließlich in der Leistungsdomäne bedeutsame Leistungseinbußen zeigen, in welcher das gängige Geschlechterstereotyp gegen ihren Erfolg sprach (Deutsch), wohingegen die Leistung im mathematischen Bereich als maskulin konnotierte Domäne unbeeinflusst blieb. Mädchen hingegen profitierten von der Darstellung generell schulversagender Jungen ebenfalls nur im Fach Deutsch, also in einem Leistungsbereich, der typischerweise einen Leistungsvorteil von Mädchen und Frauen unterstellt. Diese Ergebnisse machen deutlich, dass trotz des generellen, fachübergreifenden Bedrohungsreizes die Leistung nicht kongruent generell beeinträchtigt wird, sondern dass von einer Verstärkung (i.S.v. einer leichteren Zugänglichkeit) der geltenden fachbezogenen Geschlechterstereotype auszugehen ist, sodass Effekte durch die Stereotypenbedrohung von Jungen nur in dem Fach aufzufinden sind, in welchem Jungen auch tatsächlich eine schlechtere Leistung erbringen. Offensichtlich war die applizierte Stereotypenbedrohung trotz ihrer allgemeinen Natur im Kontrast zu den Ergebnissen aus *Studie 1* im Leistungskontext nicht in der Lage, fächerspezifische Vor- und Nachteile in der Leistungsfähigkeit von Jungen zu überlagern, da diese u.U. derart tiefgreifend internalisiert sind, dass infolge der Betonung von Geschlecht und einer darauf bezogenen Bedrohung von einer automatischen Aktivierung im konkreten Leistungskontext ausgegangen werden kann (Banaji & Hardin, 1994; Kiefer & Sekaquaptewa, 2007; Zembre et al., 2000). Geschlechterstereotype scheinen die real existierenden Unterschiede besser zu reflektieren, als das medial transportierte Stereotyp über schulversagende Jungen im Allgemeinen, sodass die mediale Darstellung statt einer generellen Leistungsbeeinträchtigung vielmehr eine Verstärkung von geschlechterstereotypisierter Leistungsunterschiede zu Folge hat: Jungen und Mädchen, welche der übermäßig negativen Darstellung von Jungen ausgesetzt waren, produzierten ein wesentlich stärker durch Geschlechterstereotype geprägtes Befundmuster, als Teilnehmer/-innen der Untersuchungsbedingung ohne Stereotypenbedrohung. In Verbindung mit den Ergebnissen aus *Studie 1* scheint es, dass männliche Schüler als Mitglieder einer stigmatisierten Gruppe betrachtet werden können, d.h. einer sozialen Kategorie zugeordnet werden, über die andere negative Erwartungen oder Stereotypen vorhalten und diese allgemeine Kategorisierung hingegen eine spezifische, nachteilige Wirkung nur auf derartige Leistungsbereiche begünstigt, welche typischerweise nicht als männliche Leistungsdomäne gelten. Diese Studie ergänzt ferner den aktuellen Forschungsstand zu den Auswirkungen einer Stereotypenbedrohung dahingehend, dass nicht wie für die meisten Studien zu dieser Thematik üblich ausschließlich ein Leis-

tungsbereich untersucht wurde (nämlich der, in welcher das Leistungsstereotyp gegen den Erfolg der stereotypisierten Gruppe spricht), sondern die Auswirkung einer generellen Stereotypenbedrohung auf zwei verschiedenen Domänen betrachtet wird, in welcher die geltenden Geschlechterstereotype sowohl *für* als auch *gegen* das positive Abschneiden der durch ein generelles Stereotyp (die allgemein negative Darstellung von Jungen in den Medien) bedrohten sozialen Gruppe der Jungen sprechen. Abschließend ermöglicht es diese Studie im Vergleich zu anderen Befunden zur Wirkung einer Stereotypenbedrohung auf Jungen in der Schule, erste Vermutungen über den Entwicklungsverlauf der Konsequenzen von *stereotype threat* zu formulieren. Während in einer vergleichsweise jungen Stichprobe auf eine generelle Bedrohung noch mit einem fächerübergreifenden Leistungsabfall zu rechnen ist (Hartley & Sutton, 2013), herrscht in der hier präsentierten Studie eine differenziertere Reaktion auf eine generelle Bedrohung vor, was ein erster Hinweis darauf sein kann, dass sich mit zunehmendem Alter stärker ausdifferenzierte Geschlechterstereotype (Eckes, 2008; Zemore et al., 2000) und fachspezifische Selbstkonzepte (Marsh & Ayotte, 2003) u.U. auf die Reaktion ggü. einer Stereotypenbedrohung dahingehend auswirken können, dass mit zunehmenden Alter keine allgemeine, sondern vielmehr eine (fach-)spezifischere Reaktion wahrscheinlicher wird. Dies lässt die Notwendigkeit erkennen, dass in Untersuchungen zu *stereotype threat* vermehrt der individuelle Entwicklungsstand einhergehend mit dem Grad der Ausdifferenzierung von Stereotypen und (fachbezogenem) Selbstkonzept einbezogen werden sollte, um die Konsequenzen einer Stereotypenbedrohung besser abschätzen zu können. Es scheint möglich, dass im Entwicklungsverlauf eine zunehmende Ausdifferenzierung von Selbstkonzept und Stereotypen-Inhalten moderiert, wie vulnerabel ein Individuum gegenüber einer generellen Stereotypenbedrohung mit genereller vs. domänenspezifischer Leistungseinbuße reagiert. Dies wird durch die These unterstützt, dass im in einem früheren Entwicklungsstadium aufgrund der vergleichsweise geringeren Differenzierung von Selbstkonzept und Stereotypen-Inhalten einfacher ist, an die Richtigkeit eines generellen Stereotyps zu glauben, als es in einem Stadium erhöhter Ausdifferenzierung dieser Komponenten möglich scheint, was eine unterstützende Bedingung für das Gelingen der Induktion einer Stereotypenbedrohung darstellen kann (Schmader et al., 2004). Der Zusammenhang zwischen dem Grad der Entwicklung des Selbstkonzeptes der untersuchten Schüler/-innen sowie der inhaltlichen Ausdifferenzierung ihrer Stereotype mit dem Ausmaß an Vulnerabilität ggü. *stereotype threat* kann zwar mit den vorliegenden Daten nicht genauer bestimmt werden, sollte aber im Hinblick auf die divergierenden Forschungsergebnisse zwischen Kindes- und Jugendalter (Hartley & Sutton, 2013; Latsch & Hannover, 2014) in zukünftigen Untersuchungen aufgrund der vorliegenden Ergebnisse verstärkt Berücksichtigung erfahren.

Zusätzlich bleibt festzuhalten, dass zur Erklärung der überdauernden Geschlechterdisparitäten (im Kontrast zu einer vergleichsweise kurzen Wirkungsdauer von *stereotype threat*) im

Bildungswesen auch fächerübergreifende Kompetenzen einen Beitrag leisten (Duckworth & Seligman, 2006; Kessels & Steinmayr, 2013), welche sich wiederum auf einen anderen Indikator von Schulleistungen, nämlich die Benotung auswirken können. In der vorliegenden Untersuchung werden derartige fächerübergreifende Kompetenzen nicht näher berücksichtigt, wenngleich die Ergebnisse der zitierten Studien darauf hinweisen, dass mangelnde überfachliche Kompetenzen von Jungen zur Erklärung einer schlechteren Benotung beitragen können (Kuhl & Hannover, 2012). Daraus kann die Vermutung abgeleitet werden, dass entsprechend der These zu Schule als Ort der Inszenierung von Geschlechtsidentität (Hannover & Kessels, 2011) überfachliche Kompetenzen im schulischen Kontext eher mit einem femininen Geschlechterselbstkonzept assoziiert werden und damit einhergehend in Konflikt zu einer maskulinen Geschlechtsidentität stehen. Demzufolge erscheint es nach Hannover & Kessels (2011) möglich, dass Jungen weniger überfachliche Kompetenzen erlernen und einsetzen, um somit ihre eigene (maskuline) Geschlechtsidentität zu stabilisieren, was allerdings mit schulischen Nachteilen verbunden ist. Wenngleich dieser Wirkungspfad noch einer empirischen Prüfung zuzuführen ist, liegt hierin ein erster Erklärungsansatz, wie sich schulische Nachteile von Jungen über eine situative Stereotypenbedrohung hinaus ausweiten und verfestigen, was ferner die Notwendigkeit eines verstärkten Einbezugs überfachlicher Kompetenzen für die weitere Forschung zu Geschlechterungleichheiten im Bildungswesen unterstreicht.

Eine weitere Möglichkeit, die Auswirkungen einer situationalen Stereotypenbedrohung dauerhafter zu gestalten besteht neben der unwahrscheinlichen Annahme, dass in schulischen Leistungskontexten stets hinreichend auslösende Reize für eine Stereotypenbedrohung durch Lehrpersonen oder Kontextfaktoren appliziert werden darin, dass durch die wahrgenommene Stereotypenbedrohung nicht nur die aktuelle fachliche Leistung beeinträchtigt wird, sondern auch damit in Zusammenhang stehende motivationale Zielorientierungen beeinflusst werden. Diese könnten wiederum zu einer Erklärung beitragen, warum Leistungsunterschiede zwischen Jungen und Mädchen generell oder aber in bestimmten Schulfächern über eine kurzfristige Stereotypenbedrohung hinaus persistieren. Eine empirische Überprüfung dieser These fand in den *Studien 3 und 4* statt. Ausgehend von Modellen zu Auswirkungen einer Stereotypenbedrohung auf motivationale Variablen (Smith, 2004, 2006, Smith et al., 2007, Thoman et al., 2013) wurde angenommen, dass durch *stereotype threat* eine Vermeidungs-Leistungszielorientierung (Elliot, 1999) begünstigt wird, welche nicht leistungsförderlich wirkt (Utman, 1997; Huang, 2012). Die Ergebnisse zeigten erneut einen differenziellen Einfluss des Treatments auf die motivationale Zielorientierung von Jungen, allerdings nicht in der erwarteten Richtung. Während in Deutsch – einem typischen „Mädchenfach“ – keinerlei Auswirkungen des Treatments auf die Zielorientierung nachgewiesen werden konnten (insbesondere keine verstärkte Einnahme einer Vermeidungs-Leistungszielorientierung vorlag), finden sich im „Jungenfach“ Mathematik bedeutsame

Unterschiede im Sinne einer verstärkten Ausrichtung auf eine lernförderliche Lernzielorientierung bei Jungen der Experimentalgruppe. Dieser Befund unterstreicht wiederum die bereits in *Studie 2* angenommene Verstärkung von Geschlechterstereotypen durch einen generellen, fächerübergreifenden Bedrohungsreiz und macht deutlich, dass auch auf diesen motivationalen Variablen ebenfalls keine generelle Reaktion auf den generellen Bedrohungsreiz folgt, wie auch in *Studie 2* mit Leistungsdaten gezeigt werden konnte. Diese Ergebnisse können dahingehend interpretiert werden, dass Jungen die motivationalen Zielorientierungen und die verstärkte Ausrichtung der lernförderlichen Komponente von Zielorientierungen auf eine Domäne, in denen die gängigen Geschlechterstereotypen für ihre Leistungsfähigkeit sprechen dazu nutzen, der übermäßig negativen Beschreibung ihrer eigenen sozialen Gruppe im Sinne einer *stereotype reactance* (Gupta et al. 2008; Kray et al., 2001, 2004; Seibt & Förster, 2004) entgegenzuwirken. Konsistent mit dieser Interpretation finden sich keinerlei Auswirkungen des experimentellen Treatments auf die Angaben zur motivationalen Zielorientierung von Mädchen in Mathematik. Da diese soziale Gruppe als generell überlegen dargestellt wurde, konnte keine Reaktanzreaktion bei Mädchen ausgelöst werden. Während in bisherigen Forschungsarbeiten zu Auswirkungen einer Stereotypenbedrohung auf Zielorientierungen nahezu ausschließlich die Komponente der Vermeidungs-Leistungszielorientierung erfasst wurde (z.B. Smith et al., 2007) ergänzt die vorliegende Arbeit somit um einen Befund, dass auch andere Bestandteile von Zielorientierungen vulnerabel auf eine Stereotypenbedrohung reagieren können. Weiterhin erscheint es möglich, dass Jungen in der Ausprägung einer erhöhten Lernzielorientierung für Mathematik nicht nur den Versuch unternehmen, dem generell negativen Bild ihrer sozialen Gruppe zu widersprechen, sondern auch dass u.U. im Rahmen der Reaktanz gegen das generelle Stereotyp das Bestreben verstärkt wird, sich tatsächlich mehr mit stereotypenkonformen (= mathematischen) Aufgaben zu beschäftigen. Seibt & Förster (2004) weisen nach, dass positive vs. negative Leistungsstereotype verschiedene regulatorische Foki, *prevention vs. promotion focus*, auslösen können. Im Lichte der vorliegenden Ergebnisse kann festgehalten werden, dass festgestellte Unterschiede in den Zielorientierungen in Mathematik zwischen Jungen der Experimental- und Kontrollgruppe mit einem *prevention focus*, einhergehend mit Annäherungsstrategien und Risikobereitschaft in der Experimentalgruppe vergleichbar sind. Damit ist ferner verbunden, dass Jungen der Experimentalgruppe weniger häufig Leistungszielorientierungen präferieren, was wiederum einem *promotion focus* von Risiko- und Fehlervermeidung gleichkommt, woraus sich schlussfolgern lässt, dass neben positiven Stereotypen (Seibt & Förster, 2004) auch Reaktanz gegen ein negatives Stereotyp den regulatorischen Fokus beeinflussen kann. Zusätzlich konnten die Ergebnisse die Erkenntnisse bisheriger Untersuchungen zu motivationalen Auswirkungen einer Stereotypenbedrohung dahingehend erweitern, als dass gezeigt werden konnte, dass nicht nur eine lernhinderliche Zielorientierung in einer negativ stereotypisierten Leistungsdomäne als Konsequenz erwartet werden kann, sondern auch in einer positiv stereotypisierten Leistungs-

domäne lernförderliche Zielorientierungen im Sinne einer *stereotype reactance* durch einen generellen Bedrohungsreiz ausgelöst werden können, die zur Intensivierung bestehender Geschlechterstereotype und einer Verfestigung dieser beitragen. Dies macht deutlich, dass es im Hinblick auf zukünftige Forschungsarbeiten wertvoll erscheint, sich nicht ausschließlich auf Leistungsdomänen zu beschränken, in welchen ein negatives Stereotyp vorherrschend ist.

Die Ergebnisse dieser ersten Studien machen deutlich, dass transportiert durch das mediale Stereotyp des männlichen Bildungsverlierers keine generelle Beeinträchtigung der Leistungsfähigkeit der von diesem Stereotyp bedrohten sozialen Gruppe stattfindet, sondern vielmehr eine Verstärkung der fachbezogenen Geschlechterstereotypen im Sinne einer signifikanten Leistungseinbuße in einer typischerweise weiblich konnotierten Domäne und einer motivationalen Zuwendung (Einnahme einer Lernzielorientierung) in einer männlich konnotierten Leistungsdomäne vorzufinden war. Bisherige Befunde können somit dahingehend erweitert werden, dass der Einbezug von sowohl positiv als auch negativ stereotypisierten Leistungsdomänen bei der Untersuchung von *stereotype threat* zu einer differenzierteren Betrachtung der leistungs- und motivationsbezogenen Konsequenzen beitragen kann.

An diese ersten Ergebnisse schließt sich die Frage an, wie Jungen auf die wahrgenommenen Veränderungen der Intergruppenkonstellation (Mädchen als überlegene Gruppe) reagieren, um der eigenen sozialen Gruppe wieder einen Überlegenheitsstatus zuweisen zu können (Wiederherstellung positiver Distinktheit; Tajfel & Turner, 1986). Die Theorie der sozialen Identität geht davon aus, dass es ein zentrales Bedürfnis eines Individuums in einer Gruppe darstellt, die eigene Gruppe in eine positive Richtung abgegrenzt von anderen sozialen Gruppen wahrzunehmen (vgl. Jetten, Spears & Manstead, 1999). Fällt der soziale Vergleich zuungunsten der eigenen Gruppe aus, postuliert die Theorie verschiedenartige Strategien sozialen Wettbewerbs (z.B. Abwertung der Fremdgruppe) oder der sozialen Kreativität (z.B. Wechsel der Vergleichsdimension), mit denen der Überlegenheitsstatus der Eigengruppe wiederhergestellt werden kann (für eine ausführlichere Darstellung siehe Turner & Reynolds, 2001). Aufbauend auf diesen theoretischen Grundannahmen wurde in drei weiteren experimentellen Studien überprüft, inwiefern Jungen Strategien des sozialen Wettbewerbs und der sozialen Kreativität einsetzen, um wieder eine positive Distinktheit der Eigengruppe im Vergleich zur vermeintlich überlegenen Fremdgruppe der Mädchen zu erreichen. Da sich nicht jede im Rahmen der Theorie vorgeschlagene Strategie zur Wiederherstellung positiver Distinktheit gleichermaßen auf die Fragestellung der vorliegenden Arbeit anwenden lässt (z.B. ist ein Wechsel der sozialen Gruppe – von der Gruppe der Jungen zur Gruppe der Mädchen – nicht möglich), beschränken sich die untersuchten Strategien auf a) direkten sozialen Wettbewerb im Sinne einer Eigengruppenbevorzugung und einer Abwertung der Fremdgruppe (*Studien 5 und 6*), Wechsel der identitätsrelevanten, sozialen Vergleichsdimension (*Studie 7*) und den Wechsel der Vergleichsgruppe für den sozialen Vergleich

(Studie 8). Die Ergebnisse aus Studien 5 und 6 zeigen keinerlei treatmentbezogene Effekte zur Herstellung positiver Distinktheit nach einer Bedrohung der sozialen Identität von Jungen. Zwar fand sich Evidenz, dass die Eigengruppe sowohl bei der Beurteilung einer stark männlichkeitsinszenierenden Vignette als auch bei der Vergabe monetärer Ressourcen von Jungen bevorzugt wurde, allerdings fiel dieser Befund unbeeinflusst vom experimentellen Treatment in der Gruppe der Jungen aus. Lediglich bei der Abwertung der Fremdgruppe konnte ein differenzieller Einfluss des experimentellen Treatments nachgewiesen werden. Nach der Bearbeitung von Material, aus dem der schlechtere Stand von Jungen in der Schule im Vergleich zu Mädchen deutlich wurde, wiesen Jungen der Experimentalbedingung deutlich stärker ausgeprägte hostilesexistische und damit Weiblichkeit abwertende Einstellungen im Vergleich zu Jungen der Kontrollgruppe auf. Darüber hinaus stimmten Mädchen der Experimentalgruppe eher benevolentsexistischen Aussagen deutlicher zu, als Mädchen der Kontrollgruppe. Es kann somit festgehalten werden, dass Jungen unter Einfluss des experimentellen Treatments Weiblichkeit stärker abwerteten, als nach der Bearbeitung einer Aufgabe mit geschlechterneutralem Inhalt und darin der Versuch gesehen werden kann, die vermeintlich statushöhere Gruppe wieder „auf ihren Platz zu verweisen“ sowie darüber positive Distinktheit herzustellen. Bei Mädchen bleibt zu diskutieren und im Zuge zukünftiger Forschung zu spezifizieren, inwiefern die stärkere Zustimmung zur benevolenten Komponente des ambivalenten Sexismus dazu genutzt wird, feindselig sexistische Einstellungen ggü. Jungen auszudrücken (Verübeln von Paternalismus [Glick & Hilt, 2000] als Ausdruck von Ärger der statusniedrigeren Gruppe über die höhere strukturelle Macht der statushöheren Gruppe), oder aber durch die erst stattfindende Entwicklung benevolentsexistischer Konzepte im Jugendalter (Glick & Hilt, 2000) bei den befragten Schülerinnen die eigentliche negative, sexistische Komponente der benevolenten Sexismus-Skala noch nicht als solche wahrgenommen wird. Dies könnte dazu führen, dass als Reaktion auf das Stereotyp des männlichen Schulversagers die berichtete höhere Zustimmung zum benevolenten Sexismus durch wahrgenommene Vorteile aus dieser Komponente des ambivalenten Sexismus getriggert wird und noch kein Problembewusstsein über die eigentlichen negativen Konsequenzen (z.B. Dardenne, Dumont & Bollier, 2007; Barreto & Ellemers, 2005; Becker & Wright, 2011) in diesem Entwicklungsstadium vorhanden ist. Für eine weitere Klärung dieses Befundes wäre es hilfreich, zukünftig auch ambivalenten Sexismus gegenüber Jungen/Männern in der Datenerhebung mit zu berücksichtigen.

Diese Ergebnisse können u.U. auf eine graduelle Abstufung der Reaktion auf die Bedrohung durch das Stereotyp des schulversagenden Jungen im Hinblick auf den Grad an Feindseligkeit, aber auch hinsichtlich der Langfristigkeit der Konsequenzen hinweisen. Eine Eigengruppenfavorisierung findet bei Jungen im Allgemeinen – ohne differenziellen Einfluss der experimentellen Manipulation – statt, jedoch erst auf einem sehr direkten Maß mit langfristigen (ge-

samtgesellschaftlichen) Konsequenzen, dem ambivalenten Sexismus, zeigt sich ein differenzieller Einfluss des Treatments, bei einer kurzfristigen Reaktion (z.B. durch die einmalige verminderte Zuweisung von Ressourcen) hingegen nicht. Es kann allerdings kritisch diskutiert werden, dass Jungen dieser Altersstufe evtl. generell stark darauf fokussiert sind, die eigene Geschlechtergruppe abzugrenzen (verminderte Permeabilität der Gruppengrenzen, Gavin & Furman, 1989) und aufzuwerten (Tarrant, 2002), was wiederum differenzielle Effekte des experimentellen Treatments erschwert, sodass diese nur im Rahmen einer extrem starken Reaktion (Sexismus) auf eine Stereotypenbedrohung aufgefunden werden können (ferner sind methodische Probleme bei der Operationalisierung der abhängigen Variablen nicht auszuschließen, vgl. Abs. 9.3.).

Aufbauend auf diese ersten Ergebnisse wurden zwei weitere Strategien als Möglichkeit der Wiederherstellung positiver Distinktheit überprüft, welche aber keinen direkten sozialen Wettbewerb, sondern die indirektere Form sozialer Kreativität geprüft haben. Hierzu wurde untersucht, ob Jungen die identitätsrelevante Vergleichsdimension Schule in Richtung außerschulischer Vergleichsdimensionen wechseln, um wieder einen Überlegenheitsstatus einzunehmen (*Studie 7*) oder aber die Vergleichsgruppe wechseln und sich eher mit einer unterlegenen Statusgruppe vergleichen (*Studie 8*). Die Ergebnisse zur ersten Fragestellung verdeutlichten, dass Jungen unabhängig von der experimentellen Beeinflussung dazu geneigt sind, außerschulische Bereiche (wie Sport, Beliebtheit bei anderen und außerschulisches Engagement) für die Vergabe einer attraktiven Position stärker zu berücksichtigen und im Rahmen der vorgelegten Rangplataufgabe auf einen niedrigeren (= wichtigeren) Rangplatz zu nominieren, als Mädchen. Wird die Relevanz verschiedener schulischer, außerschulischer und sportbezogener Eigenschaften in einem geschlossenen Antwortformat abgefragt zeigte sich, dass Jungen die Relevanz aller drei Teilbereiche unabhängig von der Untersuchungsbedingung als wichtiger beurteilten als Mädchen. Lediglich auf der Dimension Sport und Sportlichkeit konnte bei Jungen ein tendenzieller Einfluss des experimentellen Treatments dahingehend gefunden werden, dass Jungen der Experimentalgruppe diesem Leistungsbereich eine noch höhere Relevanz zuschrieben.

Zum Wechsel der Vergleichsdimensionen (*Studie 8*) machte eine initiale binäre Regressionsanalyse zwar deutlich, dass Jungen der Experimentalgruppe häufiger bereit sind abwärtsgerichtete Vergleiche anzustellen, die vertiefenden Analysen zeigten jedoch, dass dieses Ergebnismuster primär von den Reaktionen der Mädchen auf das experimentelle Treatment gespeist wurde: Mädchen der Experimentalgruppe waren zu einem großen Teil weniger häufig bereit, einen abwärtsgerichteten Vergleich mit einer weniger leistungsstarken sozialen Gruppe anzustellen, als vergleichsweise Mädchen der Kontrollgruppe. Zusätzlich bleibt festzuhalten, dass sowohl Jungen als auch Mädchen beider Untersuchungsgruppen am ehesten den Vergleich mit dem eigenen Geschlecht und der eigenen Schulform präferierten. Damit einhergehend konnte in

beiden Studien kein Hinweis gefunden werden, dass die Bedrohung der sozialen Identität durch das experimentelle Treatment Strategien der sozialen Kreativität zur Wiederherstellung positiver Distinktheit bei Jungen auslöst, wobei allerdings ein möglicher Einfluss durch eine problembehaftete Operationalisierung der abhängigen Variablen (vgl. Abs. 10.4. und 11.4.) unbedingt zu berücksichtigen bleibt.

Die Interpretationsmöglichkeiten dieser Ergebnisse werden zzgl. durch die uneindeutige Befundlage und das Fehlen zur weiteren Einordnung der Ergebnisse notwendigen empirischen Studien, die sich auf das Jugendalter fokussiert haben, deutlich erschwert. Es erscheint einerseits möglich, dass Jungen auf diesem Stand ihrer sozialen Entwicklung Strategien sozialer Kreativität keinen bedeutsamen Einfluss für das Management der Position ihrer sozialen Gruppe beimessen und u.U. das Wissen über andere sozialen Gruppen und deren Beziehungen zueinander sowie alternativer (nicht-schulischer) Vergleichsdimensionen noch nicht umfangreich genug vorhanden und ausdifferenziert ist (Adams, 1983). Dies *kann* zur Folge haben, dass die Auswirkungen sozialer Kreativität auf die Positionierung der Eigengruppe nicht hinreichend genau abgeschätzt werden und daher derartige Strategien im Jugendalter möglicherweise nicht als adäquate Mittel zur relativen Aufwertung der Eigengruppe angesehen werden. Zusätzlich stellen hostile Verhaltensweisen in dieser Phase des Jugendalters besonders häufig anzutreffende Reaktionen dar (Lindeman, Harakka & Keltikangas-Järvinen, 1997), was das Auftreten sozialer Kreativität zusätzlich unwahrscheinlicher macht. Zusätzlich weisen Gibbons & Buunk (1999) darauf hin, dass die soziale Vergleichsorientierung über verschiedene Altersspannen variabel ausfällt (vgl. auch Blanchard-Fields & Hertzog, 2000). Während Jackson et al. (1996) in einer Stichprobe von Studienanfänger/-innen nachweisen, dass in dieser Altersspanne z.T. Strategien der sozialen Kreativität eingesetzt werden, mangelt es derzeit an Studien zu früheren Altersspannen. Somit können die zitierten Befunde zwar als erste Hinweise für die o.g. Vermutung angesehen werden, allerdings bleibt diese Argumentation in weiteren Studien, die sich auf derartige Strategien im Jugendalter fokussieren und den methodischen Limitationen der vorliegenden Studien begegnen (vgl. Abs. 10.4. und 11.4.) noch genauer empirisch zu überprüfen. Im Kontrast dazu ist denkbar, dass die Gruppe der Jungen auch außerschulisch nicht (mehr) als deutlich überlegen wahrgenommen wird und daher ein Wechsel der Vergleichsdimension hin zu außerschulischen Dimensionen nicht hinreichend erfolgversprechend ausfällt. Für einige außerschulische Teilbereiche ist zwar ein Überlegenheitsstatus von Jungen empirisch nachgewiesen (alltägliches Problemlösen: OECD, 2014; Sport: Eccles & Harold, 1991; Schmalz & Kerstetter, 2006) oder zumindest Inhalt gängiger Geschlechterstereotype (Führungsverhalten: Diekman & Goodfriend, 2006; Umgang mit neuen Medien: Papastergiou, 2008; für einen allgemeinen Überblick siehe Kite, Deaux & Haines, 2008). Allerdings scheinen diese angenommenen oder tatsächlichen außerschulischen Vorteile durch Jungen nicht umfassend reflektiert zu werden (vgl. Ergebnisse *Studie 1*). Die dort

antizipierte, gesellschaftliche Wahrnehmung von Jungen deutete darauf hin, dass Mädchen außerschulisch sogar positiver charakterisiert wurden, als Jungen. Trotz der Einschränkung, dass hier eine Metaperspektive einzunehmen und nicht die eigene Meinung wiederzugeben war, bleibt zu diskutieren inwiefern der Wechsel der sozialen Vergleichsdimension für Jungen dieser Altersgruppe somit eine ernstzunehmende Alternative für die Wiederherstellung eines Überlegenheitsstatus darstellen kann, da über Mädchen auch im außerschulischen Bereich ein positiveres Bild antizipiert wird. Es kann somit nicht eindeutig geklärt werden, ob trotz der empirisch nachweisbaren oder in den Geschlechterstereotypen enthaltenen Vorteile für Jungen in einigen außerschulischen Domänen die Antizipation von ebenfalls außerschulischen Vorteilen für Mädchen dazu führt, dass Jungen einen Wechsel der Vergleichsdimension als nicht geeignet zur Verbesserung des relativen Status der Eigengruppe wahrnehmen. Für zukünftige Forschungsarbeiten lässt sich hieraus ableiten, dass es günstig erscheint, *a priori* Fähigkeits- und Leistungsbereiche zu identifizieren, in denen die Überlegenheit von Jungen nicht nur empirisch angenommen oder nachgewiesen ist, sondern dies von der entsprechenden sozialen Gruppe auch so wahrgenommen wird. Speziell in Bezug auf die zweite untersuchte Strategie sozialer Kreativität (soziale Abwärtsvergleiche) kann subsummiert werden, dass zahlreiche Studien im Erwachsenenalter (Buunk et al., 1990; Suls et al., 2002; Wheeler & Myake, 1992) nachweisen können, dass soziale Vergleiche zur Wiederherstellung positiver Distinktheit eingesetzt werden, wohingegen die Praktikabilität dieser Strategie im Jugendalter ebenfalls bisher nicht umfassend untersucht ist. Die Abweichung der vorliegenden Ergebnisse zur o.g. Befundlage kann möglicherweise damit erklärt werden, dass Abwärtsvergleiche über eine andere Schulform möglicherweise zu bedrohlich wirken, als dass diese Form Wiederherstellung positiver Distinktheit geeignet erscheint, eine bedrohte soziale Identität zu stabilisieren. Dafür spricht ferner der Nachweis, dass unterlegene Gruppen eher Intragruppenvergleiche präferieren, als Intergruppenvergleiche anzustellen (Lockwood, 2002; Zagefka & Brown, 2005) und die Tatsache, dass Jungen (*aber* auch Mädchen) den Vergleich mit dem eigenen Geschlecht und der eignen Schulform am stärksten präferieren.

Abschließend kommt dem Faktor Geschlecht im Jugendalter eine herausragende Bedeutung (erhöhte Salienz) zu, darüber hinaus ist für Jungen die eigene Gruppenmitgliedschaft von größerer Bedeutung, als dies für Mädchen der Fall ist (Cross & Madson, 1997). Dies kann trotz ausbleibender Bedrohung auf Grundlage der Geschlechtergruppe die Wahrscheinlichkeit für Eigengruppenfavorisierungen und das Bedürfnis nach einer positiven Abgrenzung der *Ingroup* auch in der Kontrollgruppe erhöhen. Dafür sprechen Hinweise von Tarrant (2002), welcher feststellen konnte, dass im Jugendalter der eigenen Geschlechtergruppe generell vermehrt positive Eigenschaften zugeschrieben werden, d.h. der *ingroup bias* im Allgemeinen sehr stark ausgeprägt ist. Insgesamt spricht die Gesamtheit der Befunde aus den Studien zur Theorie der sozialen Identität für die Annahme, dass eher hostile Strategien eingesetzt werden und sozialer Krea-

tivität keine besondere Rolle zukommt, was zum Großteil mit den Befunden von Lindeman et al. (1997) übereinstimmt. Biernat, Vescio & Green (1996) liefern darüber hinaus Befunde zu einer weiteren Strategie sozialer Kreativität im Sinne einer Aufwertung der Eigengruppe. Als Ausblick einer weiteren noch auf das Jugendalter zu übertragende Strategie sozialer Kreativität kann hieraus abgeleitet werden, dass eine verstärkte positive Selbststereotypisierung als Reaktion auf die soziale Bedrohung der Eigengruppe als protektive und sozial-kreative Strategie eingesetzt werden kann, um eine positive Distinktheit der Eigengruppe zu bewahren oder wiederherzustellen. Diese Strategie stellt eine besondere und weniger hostile Form der Aufwertung der Eigengruppe dar, welche nicht im Rahmen der vorliegenden Arbeit geprüft wurde. Weitere Befunde, die Strategien der sozialen Kreativität nachweisen konnten, beziehen sich auf einen höheren Altersbereich (z.B. Jackson et al., 1996; Wheeler & Myake, 1992) oder den organisationalen Kontext (z.B. Terry & O'Brien, 2001) und konnten in der vorliegenden Untersuchung nicht auf das Jugendalter übertragen werden. Dies macht die Notwendigkeit weiterer Forschung zu sozialen Identitäten im Jugendalter deutlich. Ferner bleiben mögliche Moderatorvariablen wie u.a. die Identifizierung mit der Eigengruppe (Eckert & Imhoff, 2013), die soziale Dominanzorientierung, welche die Wahrscheinlichkeit hostiler Reaktionen auf eine Bedrohung der sozialen Identität erhöhen kann (Sibley, Wilson & Duckitt, 2007) und die Unterscheidung zwischen biologischem und psychologischem Geschlecht (Deaux, 1985) verstärkt in zukünftige Forschungsarbeiten einzubeziehen. Zwar liefern die vorliegenden Ergebnisse erste Hinweise, allerdings sind diese nur mit Vorsicht zu interpretieren – zu einer abschließenden Klärung bedarf es weiterer Forschungsanstrengungen, um die gefundenen Ergebnisse zu replizieren oder zu spezifizieren.

Reflexionen für Wissenschaft und Praxis

Wie bereits im Rahmen der Gesamtdiskussion festgehalten wurde, existieren derzeit nur wenige Erkenntnisse zur Vulnerabilität von Jungen gegenüber einer Stereotypenbedrohung, deren motivationalen Auswirkungen und darauf aufbauend zur Wiederherstellung und Absicherung einer positiven sozialen Identität von Jungen im Schulkontext eingesetzter Verhaltensstrategien. Während im (jungen) Erwachsenenalter der Zusammenhang von *stereotype threat* und der Mathematikleistung für Frauen häufig untersucht und nachgewiesen wurde (Nguyen & Ryan, 2008) liefern Erkenntnisse zu einer allgemeinen Bedrohung von Jungen (Hartley & Sutton, 2013) oder *stereotype threat* in sprachlichen Domänen (Eckert & Imhoff, 2013) divergierende Ergebnisse. Für die sozialpsychologische Forschung leitet sich aus den Ergebnissen der vorliegenden Arbeit weiterer Forschungsbedarf ab, um die Auswirkungen von Stereotypenbedrohungen auf Jungen auch im Hinblick auf ihre langfristigen Auswirkungen genauer zu untersuchen. Einerseits erscheint es für die Klärung der Frage, ob Stereotypenbedrohungen derart langfristige Konsequenzen nach sich ziehen können, dass hieraus die eingangs beschriebenen überdauernden Nachteile für Jungen im Kontext Schule wertvoll, längsschnittliche Untersuchungen anzu-

stellen, um Veränderungen durch *stereotype threat* über eine längere Zeitspanne beurteilen zu können. Weiterhin konnte nachgewiesen werden, dass zahlreiche Drittvariablen (wie Identifizierung mit der Eigengruppe, Identifizierung mit der Domäne usw.) einen Einfluss auf die Art und Dauerhaftigkeit der Konsequenzen einer Stereotypenbedrohung haben, woraus sich ein verstärkter Einbezug potentiell moderierender oder mediierender Variablen in zukünftigen Forschungsvorhaben zu diesem Thema ableiten lässt, wenngleich in zusammenfassenden Forschungsarbeiten die Notwendigkeit betont wird, dass experimentelle Untersuchungen zu Stereotypenbedrohungen möglichst nahe am Originalparadigma von *stereotype threat* orientiert werden sollten (Stoet & Geary, 2012). Auch die durch eine Stereotypenbedrohung induzierte unsichere soziale Identität und die daraus resultierenden Bemühungen zur Wiederherstellung positiver Distinktheit der Eigengruppe (Tajfel & Turner, 1986) wurden bisher kaum auf das Jugendalter angewendet. Die ersten Ergebnisse dieser Arbeit können als *mögliche* Indikatoren dafür dienen, dass im Jugendalter noch nicht alle im Rahmen der Theorie postulierten Strategien als adäquat angesehen und eingesetzt werden, jedoch bleiben diese Ergebnisse durch weitere Forschung noch abzusichern. Zusammenfassend ergibt sich ein Forschungsbedarf sowohl zur *stereotype threat*-Theorie als auch zur Theorie der sozialen Identität, der bisherige Ergebnisse auf das Jugendalter anwendet und somit genauere Aussagen über die Entwicklung beider Konstrukte über die Lebensspanne erlaubt und weniger auf das (junge) Erwachsenenalter fokussiert. Ebenfalls wäre wünschenswert, die Forschungsanstrengungen um negative Auswirkungen von Stereotypenbedrohungen stärker auf das psychologische und weniger das biologische Geschlecht zu fokussieren. Im Rahmen der vorliegenden Arbeit wurde entsprechend der zugrundeliegenden theoretischen und empirischen Paradigmen nur auf das biologische Geschlecht zentriert, was der Forschungstradition um *stereotype threat* entspricht. Im Rahmen der aktuellen Debatte um „*doing gender*“ und das psychologische Geschlecht (Gildemeister, 2004) sowie der damit einhergehenden Annahme, dass Unterschiede zwischen beiden Geschlechtergruppen geringer ausfallen, als Varianzen innerhalb eines Geschlechts (Hyde, 2005) erscheint es jedoch wertvoll, in zukünftigen Untersuchungen verstärkt auch psychologische Geschlechterunterschiede in den Forschungsfokus zu rücken, um der breiten Variabilität von männlichen vs. weiblichen Identitäten, resp. maskulinen und femininen Aspekten von Selbst und Persönlichkeit stärker Rechnung zu tragen.

Für die schulische Praxis weisen die Ergebnisse der vorliegenden Arbeit darauf hin, dass *a priori* Leistungszuschreibungen auf Grundlage des biologischen Geschlechts zu vermeiden sind, da diese sowohl für Jungen als auch für Mädchen leistungshemmende und motivational wenig lernförderliche Konsequenzen haben können sowie bestehende Geschlechterunterschiede aufrecht erhalten. Darüber hinaus konnte gezeigt werden, dass durch die Beschreibung einer Geschlechtergruppe als leistungsmäßig überlegen eine unsichere soziale Identität hergestellt wurde, welche in der Hauptsache mit hostilen Gegenreaktionen (hier: der Abwertung von Weib-

lichkeit durch erhöhten hostilem Sexismus) in Verbindung zu stehen scheint. Das Auftreten derartiger Reaktionen wird dabei in dem Maße verstärkt, in dem Geschlecht im schulischen Kontext betont und dadurch hervorgehoben wird. Zum einen können somit negative Konsequenzen aus einer Stereotypenbedrohung in der Schule vermieden werden, indem die Aktivierung von Geschlecht als zentraler Kategorie und der damit verbundenen Erhöhung der Salienz von Geschlecht für die soziale Identität durch eine Pädagogik entgegengewirkt wird, in der sensibel mit geschlechterspezifischen Leistungs- und Rollenzuweisungen umgegangen wird (Stricker & Ward, 2004). Zum anderen wäre möglich, durch gezielte Aufklärung über die Wirkmechanismen und Konsequenzen von Stereotypenbedrohungen zumindest Leistungs- und motivationalen Einbußen entgegenzuwirken („*knowing is half the battle*“; Johns et al., 2005). Weitere Möglichkeiten zum Umgang mit Stereotypenbedrohungen in der Schule bestehen in einer Schulung zum verstärkten Einsatz selbstbegründenden Leistungsverhaltens (Schimmel, Arndt, Banko & Cook, 2004), der Bereitstellung von Rollenmodellen, die trotz angenommener Leistungsnachteile in einer Domäne exzellente Leistungen erzielen (Blanton, Crocker & Miller, 2000) oder eine inkrementelle Sichtweise auf Intelligenz zu etablieren (Goff, Steele & Davies, 2008). Neben dem bewussten Vermeiden der sozialen Kategorie Geschlecht in der Schule können diese Strategien zusätzlich dazu beitragen, negative affektive sowie leistungs- und motivationsbezogene Aspekte einer Stereotypenbedrohung abzumildern, vor allem auch im Hinblick darauf, dass Geschlecht im Jugendalter eine zentrale Kategorie des Selbst darstellt und somit neben einem sensiblen Umgang mit der Kategorie Geschlecht im schulischen Kontext Möglichkeiten bereitgestellt werden, einer vom Lehrkraftverhalten unabhängigen Bedrohung durch Stereotype wirksam begegnen zu können. Es bleibt anzunehmen, dass durch diese zentralen Interventionsstrategien die soziale Identität stabilisiert werden kann, sodass kein weiteres Bedürfnis nach positiver Distinktheit getriggert wird und damit einhergehend die hostile Reaktionen, insbesondere im Hinblick auf den Sexismus, ausbleiben. Sollte dies nicht der Fall sein, dann ist eine weitere Aufgabe von Schule darin zu sehen, die Entwicklung von positiven Strategien zur Wiederherstellung positiver Distinktheit zu fördern. Damit geht einher, über die negativen Konsequenzen von Sexismus als hostile Strategie zur Aufwertung der Eigengruppe aufzuklären (Becker & Swim, 2011; Kilmartin, Green, Smith, Heinzen, Kuchler & Kolar, 2008) und zu einem verstärkten Einsatz sozial-kreativer Strategien für das Management der sozialen Identität anzuleiten, welche den Ergebnissen folgend im Jugendalter bisher kaum eingesetzt werden.

Neben dem großen Einfluss, den Schule auf die Vermeidung und Abschwächung negativer Konsequenzen für den Bildungsverlauf von Jungen (und Mädchen) nehmen kann sind trotz des Umstandes, dass Schule einen zentralen Lebensmittelpunkt von Jugendlichen darstellt auch andere Kontextfaktoren mit einzubeziehen, welche schulische Bemühungen um mehr Geschlechtergerechtigkeit ergänzen sollten. Auch Einflüsse des Elternhauses determinieren nachhaltig, wie stark geschlechterstereotyp Jungen ihre Lebenswelt gestalten (Eccles, Jacob & Harold,

1990; Tenenbaum & Leaper, 2002) und welche Einstellungen sie zu Fremdgruppen aufbauen (z.B. maskuline Ideologien und Verhältnis zu Frauen; Luddy & Thompson, 1997). Eine Verbesserung der schulischen Lernumwelten im Sinne einer Vermeidung geschlechterstereotyper Zuschreibungen kann unter diesem Gesichtspunkt nur gelingen, wenn die Verminderung der Salienz von Geschlecht zumindest z.T. auch in die familiäre Lebenswelt integriert wird. Ferner kann die Medienberichterstattung, welche Jungen derzeit als die generellen Schulversager *labelt* durch eine ausgewogenere Berichterstattung, die sich an empirischen Ergebnissen orientiert dazu beitragen, dass schulisches Engagement nicht als weiblich dargestellt und rezipiert wird. Dies verhindert a) die Herausbildung einer Geschlechterkonnotation von Schule und kann maßgeblich dazu beitragen, schulisches Engagement von Jungen nicht als inkompatibel zur eigenen Identitätsentwicklung zu erleben (vgl. Kessels & Hannover, 2004, 2006) und b) minimiert die aus der medialen Darstellung resultierende Selbstwertbedrohung für Jungen. Für die Bildungsadministration und Bildungspolitik ergeben sich dringende Hinweise, das Zusammenspiel von Geschlecht und Schulerfolg differenziert in allen Phasen der Aus- und Weiterbildung von Lehrkräften wesentlich stärker zu berücksichtigen. Nicht selten stellt der Umgang mit Heterogenität im Bildungswesen nur einen vergleichsweise geringen Anteil an der erziehungswissenschaftlichen Ausbildung für Lehramtsanwärter/-innen dar. Für das Schuljahr 2013/14 stehen für Berliner Lehrkräfte insgesamt 436 breit gefächerte Fortbildungsveranstaltungen zur Verfügung, wovon allerdings keine einen bedeutsamen Bezug zu Geschlecht im schulischen Kontext und möglichen Konsequenzen für das Lern- und Arbeitsverhalten sowie den Schulerfolg von Jugendlichen zum Gegenstand hat (Senatsverwaltung für Bildung und Forschung, 2013). Auf Grundlage dieser Zahlen wäre es wünschenswert, das Fortbildungsangebot sukzessive zu erweitern, sodass die Ergebnisse bisheriger Forschungsarbeiten nicht nur im Rahmen der wissenschaftlichen Schulforschung diskutiert, sondern auch an unmittelbar Betroffene aus der Praxis weitergegeben werden. Die bedarfsgerechte Aufbereitung der bisherigen und zukünftigen Forschungsergebnisse für Lehrkräfte und ein konsequenter Einbezug in Aus- und Weiterbildung stellen somit einen zentralen Faktor für das Gelingen einer innovativen, geschlechtergerechteren Pädagogik dar. Auch ist fraglich, inwieweit eine Fokussierung von Bemühungen für mehr Geschlechtergerechtigkeit im Bildungswesen durch Kampagnen, in denen Geschlecht zu einer zentralen Kategorie hervorgehoben wird (z.B. „Girl’s Day“) neben ihren positiven Aspekten (wie Kontakt mit positiven Rollenmodellen) auch dazu beitragen, gängige Geschlechterrollenstereotype in der breiten Öffentlichkeit und auch bei Schüler/-innen selbst zu verfestigen. Geeigneter wäre u.U., durch die konsequente Verminderung der Salienz von Geschlecht im Bildungswesen eine umfassende geschlechteregale Einstellung zu erreichen, sodass Auswirkungen wie allgemeine und fachspezifische Geschlechterdisparitäten im Schulerfolg oder geschlechterkonnotierte Berufswahlen *eo ipso* positiv beeinflusst werden und somit den Weg einer insgesamt geschlechtergerechten Schule – für Mädchen wie für Jungen – ebnen können.

ZUSAMMENFASSUNG

Die schulischen Leistungen von Jungen sind in den vergangenen Jahren nicht nur deutlich stärker in den Forschungsfokus geraten, sondern wurden auch in den Massenmedien breit diskutiert (Weaver-Hightower, 2003). Zentraler Inhalt des öffentlichkeitswirksam geführten Diskurses ist, dass Jungen relativ zu Mädchen zurückgefallen sind, so dass sie heute, im Unterschied zu der Situation in früheren Kohorten, weniger hochwertige Schulabschlusszertifikate erwerben als Mädchen. Empirisch lässt sich feststellen, dass Jungen mit ansteigendem Ausbildungsniveau immer weniger vertreten sind. So besuchen die Hauptschule mehr Jungen als Mädchen, hingegen das Gymnasium mehr Mädchen als Jungen. Bezüglich akademischer Leistungen kann festgehalten werden, dass Jungen in der Lesekompetenz – einer Schlüsselkompetenz für den akademischen Erfolg – deutlich geringere Leistungen aufweisen als Mädchen, während allerdings in der Mathematik ein vergleichsweise geringerer, aber signifikanter Leistungsvorsprung zugunsten der Jungen besteht. Insofern ist das durch die Medien propagierte generelle Stereotyp über den männlichen Schulversager nicht angemessen, da sich die Kompetenzen von Mädchen und Jungen unverändert in Abhängigkeit der Geschlechtskonnotation der fachlichen Domäne voneinander unterscheiden.

Dem Forschungsansatz von Hannover & Kessels zufolge (z.B. Hannover & Kessels, 2004, 2011; Kessels & Hannover, 2004, 2006) ist Schule nicht nur ein Ort des Kompetenzerwerbs, sondern auch ein Ort, an dem Kinder und Jugendliche ihre Identität entwickeln und aushandeln. Genauer wird in Bezug auf die geschlechtsbezogene Identitätsentwicklung angenommen, dass Kinder und Jugendliche die Geschlechtskonnotation von schulischen Lern- und Interaktionsangeboten prüfen und ihr Verhalten so ausrichten, dass ihre Identität als Junge bzw. Mädchen abgestützt wird.

In dieser Sichtweise kann die mediale Darstellung von Jungen als "Verlierern in der Schule" dazu beitragen, dass Jungen sich von Schule distanzieren und in ihrer Leistungsfähigkeit durch Stereotypenbedrohung beeinträchtigt werden. Das Paradigma der Stereotypenbedrohung geht davon aus, dass durch die Befürchtung, ein negatives Stereotyp über die eigene Gruppe zu bestätigen, die Leistungsfähigkeit der betroffenen Person reduziert ist. Im Rahmen der vorliegenden Arbeit wird untersucht, inwiefern Schülerinnen und Schüler das Stereotyp des männlichen Schulversagers überhaupt wahrnehmen (Studie 1) und welche Konsequenzen aus dieser Wahrnehmung für die Leistung (Studie 2) und die Motivation (Studien 3 und 4) von Jungen in den Fächern Deutsch und Mathematik erwachsen. Ferner wird im Rahmen der vorliegenden Arbeit überprüft, inwiefern die Bedrohung des Überlegenheitsstatus von Jungen das Intergruppenverhalten im schulischen Kontext beeinflusst. Die Theorie der sozialen Identität (Tajfel & Turner, 1986) postuliert hierzu Möglichkeiten, wie sich eine durch Bedrohung induzierte, unsichere soziale Identität durch Wiederherstellung positiver Distinktheit stabilisieren lässt, so z.B. durch die Abwertung

der Fremdgruppe (Studien 5 und 6), den Wechsel identitätsrelevanter, sozialer Vergleichsdimensionen (Studie 7) oder den Wechsel der sozialen Vergleichsgruppe (Studie 8).

Die Ergebnisse der Studie 1 zeigen, dass Jugendliche Jungen im schulischen Kontext deutlich mehr negative Eigenschaften zuschreiben als Mädchen, die überwiegend durch positive schulbezogene Eigenschaften charakterisiert werden. Die Ergebnisse der Studie 2 verdeutlichen, dass sich dieses negative Stereotyp auf die Leistungsfähigkeit von Jungen im eher feminin konnotierten Leistungsbereich Deutsch, nicht jedoch in der eher maskulin wahrgenommenen Domäne der Mathematik, ungünstig auswirkt. Nach der Bearbeitung eines kurzen Lesetextes, aus welchem sich eine generell geringere schulische Leistungsfähigkeit von Jungen ableiten ließ, wurde die Leistung von Jungen in dieser Gruppe in Deutsch im Vergleich zu einer Kontrollgruppe bedeutsam gemindert, wohingegen Mädchen in der Experimentalgruppe sogar eine im Vergleich zur Kontrollgruppe gesteigerte Leistung im Fach Deutsch aufwiesen (Studie 2). Für Mathematik zeigten sich keine bedeutsamen Leistungsunterschiede als Ergebnis des experimentellen Treatments. In Studie 3 wurde der Effekt eines ähnlichen experimentellen Treatments auf die Motivation untersucht und in Studie 4 repliziert. Nach Bearbeitung einer kurzen Statistik, aus welcher das schlechtere Abschneiden von Jungen in der Schule hervorging, konnte für das Fach Deutsch keinerlei Unterschied auf der Ebene motivationaler Zielorientierungen gefunden werden, wohingegen für die Mathematik Jungen der Experimentalbedingung im Vergleich zu einer Kontrollgruppe eine stärkere Lernzielorientierung aufwiesen (Studien 3 und 4). Diese Ergebnisse sprechen dafür, dass Stereotypenbedrohung für Jungen zu einer Verstärkung der Geschlechtstypisierung in ihrer motivationalen und leistungsbezogenen Entwicklung in der Schule beiträgt. Bezüglich der Studien 5 bis 8 zur Theorie der sozialen Identität ließ sich feststellen, dass Jungen nach Bearbeitung des experimentellen Treatments tendenziell stärker dazu geneigt waren, direkten sozialen Wettbewerb zu betreiben und insbesondere Weiblichkeit abzuwerten, was sich in einer erhöhten Ausprägung hostilexistischer Einstellungen gegenüber Mädchen im Vergleich zu Jungen der Kontrollgruppe äußerte. Demgegenüber berichteten Mädchen der Experimentalbedingung mehr benevolentsexistische Einstellungen gegenüber Mädchen als Mädchen der Kontrollgruppe. In den Studien 5 bis 8 zur Wiederherstellung positiver Distinktheit konnte weder für den Wechsel der Vergleichsdimension, noch für den Wechsel der Vergleichsgruppe ein differenzieller Effekt in Abhängigkeit der experimentellen Manipulation festgehalten werden. Dies kann bedeuten, dass Jungen der untersuchten Altersstufe „soziale Kreativität“ möglicherweise nicht als adäquate Strategie zur Absicherung einer Bedrohung ihrer sozialen Identität ansehen, sondern sie verstärkt auf sozialen Wettbewerb zurückgreifen.

Die Ergebnisse werden abschließend zusammengefasst und hinsichtlich ihrer Implikationen für das Forschungsfeld der Sozial- und Pädagogischen Psychologie, für das Anwendungsfeld Schule und für Bildungsadministration und Bildungspolitik diskutiert.

SUMMARY

Over the past few years, the academic performance of boys has not only increasingly become the focus of research but has also been widely discussed in the mass media (Weaver-Hightower, 2003). The key content of this discourse, which has attracted considerable public attention, is that boys have fallen behind academically, compared to girls. In contrast to the situation in previous cohorts, they are therefore currently gaining school leaving certificates of a lower level than those gained by girls. Empirically, it is noticeable that boys are increasingly less represented as the level of education increases; i.e. in Germany, a higher proportion of boys attends the lowest level of secondary school ("Hauptschule") whilst girls are over-represented in the schools that offer the highest level of secondary school education ("Gymnasium"). In terms of academic performance, it is fair to say that as far as reading competence is concerned – which is a key skill for academic success – boys are performing considerably less well than girls, although the difference is comparatively less pronounced yet still marked in mathematics, where the boys are performing better. The stereotypical label of "male non-achiever" generally propagated by the media is therefore applied unjustly as the performance of girls and boys continues to differ in line with the gender connotations of a particular subject domain.

According to the research approach practised by Hannover & Kessels (e.g. Hannover & Kessels, 2011, Kessels & Hannover, 2004, 2006), school is not only a place for competence acquisition but also a place where children and young adults develop and negotiate their identity. Focusing in more detail on gender-based identity development, it appears that children and young adults assess the gender connotations of educational and social interaction offers at school and gear their behaviour towards supporting their male or female identity.

Considering this aspect, the media representation of boys as "academic non-achievers" can contribute to boys distancing themselves from school education. Stereotyping can therefore pose a threat which has an adverse impact on their performance at school.

The paradigm of stereotyping as a threat is based on the fear that affirming a negative stereotype through the own group reduces a particular individual's achievement potential. The extent to which male and female students in fact perceive the stereotypical "male academic non-achiever" (Study 1), and also the consequences arising from this perception in terms of the performance of boys in the subjects of German and mathematics (Study 2) and in terms of motivation (Studies 3 and 4) will be examined within the scope of this work. It will further examine to what extent this threat to the superiority status of boys impacts on intergroup behaviour in a school context. In this respect, the theory of social identity (Tajfel & Turner, 1986) postulates approaches to how an insecure social identity induced by threat can be stabilised through the reestablishment of positive distinctions, for example through degradation of the other group (Studies 5 and 6), ad-

justment of the identity relevant social comparison dimensions (Study 7) or by changing the social comparison group (Study 8).

The results of Study 1 show that in a school context, young adults ascribe considerably more negative characteristics to boys than they do to girls, who are characterised mainly by positive attitudes when it comes to school. The results of Study 2 clearly illustrate that this negative stereotyping with regard to the academic performance of boys has an adverse impact on the achievements of boys in the subject of German, which tends have feminine connotations, although it has no effect on the performance of boys in the subject of mathematics, a domain that tends to be perceived as masculine. Subsequent to being given a short reading text to work through, which suggested the conclusion that the boys in the group generally performed less well academically, their achievements in the subject of German were in fact considerably lower compared to the achievements in a control group, whereas the girls in the group that participated in the experiment even performed better than those in the control group in the subject of German (Study 2). There appeared to be no significant differences in performance in mathematics as a direct result of this experimental approach. Study 3 examined the effect of a similar experimental situation on motivation; this was replicated in Study 4. After the analysis of a brief statistic which showed that boys did worse in school, the level of motivational goal orientation appeared to be unaffected in the subject of German whereas in mathematics, the boys in the group taking part in the experiment showed a stronger orientation towards the learning aims than those in the control group (Studies 3 and 4). These results indicate that the threat of stereotyping boys contributes to increased gender typecasting in their motivation and performance related development at school. As far as Studies 5 to 8 on the theory of social identity are concerned, these showed that the boys who had taken part in the experiment tended to be more disposed towards entering into direct social competition, and in particular towards the degradation of femininity, which was expressed through more markedly hostile sexist attitudes towards girls, compared to the boys in the control group. In contrast, the girls who took part in the experiment reported an increase in benevolent sexist attitudes towards girls, compared to those in the control group. Studies 5 to 8 on the reestablishment of positive distinctions were inconclusive, as the experimental manipulation showed no differentiating effects in terms of adjustment of comparison dimension or change of comparison group. This may mean that the boys in the age group examined possibly do not consider "social creativity" to be an adequate strategy against threats to their social identity but in fact tend to rely more on social competition in this respect. By way of conclusion, the results are summarised and discussed with regard to their implications for the field of social and educational psychology research, for the application environment school and for educational administration as well as education policy.

LITERATURVERZEICHNIS

- Abele, A. (2003). The dynamics of masculine-agentive and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology, 85*, 768–776.
- Abrams, D. & Hogg, M. (1990). *Social identity theory: Constructive and critical advances*. New York, NY: Springer.
- Adams, G. (1983). Social Competence During Adolescence: Social Sensitivity, Locus of Control, Empathy and Peer Popularity. *Journal of Youth and Adolescence, 12*, 203–211.
- Ahmed, A. (2007). Group identity, social distance and intergroup bias. *Journal of Economic Psychology, 28*, 324–337.
- Aiken, L. & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. London: Sage.
- Allport, G. (1968). The historical background of modern social psychology. In G. Lindzey & E. Aronson (Hrsg.), *The Handbook of Social Psychology* (S. 1-80). Reading, MA: Addison-Wesley.
- Ambady, N., Shih, M., Kim, A. & Pittinsky, T. (2001). Stereotype Susceptibility in Children: Effects of Identity Activation on Quantitative Performance. *American Psychological Society, 12*, 385–390.
- Ammermüller, A. & Dolton, P. (2006). Pupil-teacher gender interaction effects on scholastic outcomes in England and the USA. *ZEW Discussion Papers, 60*, o.S.
- Anders, Y., McElvany, N. & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte? In K. Maaz, J., Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule – Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 313–331). Bonn: BMBF.
- Appel, M. (2008). Medienvermittelte Stereotype und Vorurteile. In B. Batinic & M. Appel (Hrsg.), *Medienpsychologie* (S. 313–336). Heidelberg: Springer.
- Aronson, J., Fried, C. & Good, C. (2002). Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence. *Journal of Experimental Social Psychology, 38*, 113–125.
- Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C. & 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.
- Artelt, C., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–111). Münster: Waxmann.
- Ashford, B. & Kreiner, G. (1999). "How can you do it?": Dirty Work and the Challenge of Constructing a Positive Identity. *Academy of Management Review, 24*, 413–434.
- Atkinson, J. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359–372.
- Banaji, M. & Hardin, C. (1994). Automatic Stereotyping. *Psychological Science, 7*, 136–141.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, M. & Dweck, C. (1985). *The relationship of conceptions of intelligence and achievement goals to achievement related cognition, affect, and behavior*. Harvard University.

- Barreto, M. & Ellemers, N. (2005). The burden of benevolent sexism: How it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology, 35*, 633–642.
- Barron, K. & Harackiewicz, J. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology, 80*, 706–722.
- Basow, S. & Rubin, L. (1999). Gender influences on adolescent development. In N. Johnson, M. Roberts & J. Worell (Hrsg.), *Beyond appearance: A new look at adolescent girls* (S. 25–52). Washington, DC: American Psychological Association.
- Baumeister, R. & Showers, C. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*, 361–383.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U. et al. (2003). *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Becker, J. & Swim, J. (2011). Seeing the Unseen: Attention to Daily Encounters With Sexism as Way to Reduce Sexist Beliefs. *Psychology of Women Quarterly, 35*, 227–242.
- Becker, J. & Wright, S. (2011). Yet Another Dark Side of Chivalry: Benevolent Sexism Undermines and Hostile Sexism Motivates Collective Action for Social Change. *Journal of Personality and Social Psychology, 101*, 62–77.
- Beilock, S., Rydell, R. & McConnell, A. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General, 136*, 256–276.
- Bem, S. (1983). Gender Schema Theory and Its Implications for Child Development: Raising Gender-Aschematic Children in a Gender-Schematic Society. *Signs, 8*, 598–616.
- Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*, 354–364.
- Ben-Zeev, T., Fein, S. & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41*, 174–181.
- Bettencourt, B., Charlton, K., Dorr, N. & Hume, D. (2001). Status differences and in-group bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin, 127*, 520–542.
- Bhana, D. (2008). 'Six packs and big muscles, and stuff like that'. Primary school-aged South African boys, black and white, on sport. *British Journal of Sociology of Education, 29*, 3–14.
- Bierhoff, H. (2006). *Sozialpsychologie: ein Lehrbuch*. Stuttgart: Kohlhammer Verlag.
- Biernat, M., Vescio, T. & Green, M. (1996). Selective Self-Stereotyping. *Journal of Personality and Social Psychology, 71*, 1194–1209.
- Billig, M. & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*, 27–52.
- Blanchard-Fields, F. & Hertzog, C. (2000). Age differences in social schematicity. In U. von Hecker, S. Dutke & G. Sedek, *Generative mental processes and cognitive resources* (S. 175–198). Dordrecht: Springer.
- Blanton, H., Buunk, B., Gibbons, F. & Kuyper, H. (1999). When Better-Than-Others Compare Upward: Choice of Comparison and Comparative Evaluation as Independent Predictors of Academic Performance. *Journal of Personality and Social Psychology, 76*, 420–430.
- Blanton, H., Crocker, J. & Miller, D. (2000). The Effects of In-Group versus Out-Group Social Comparison on Self-Esteem in the Context of a Negative Stereotype. *Journal of Experimental Social Psychology, 36*, 519–530.
- Blascovich, J., Spencer, S., Quinn, D. & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological science, 12*, 225–229.

- Blascovich, J. & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology*, 28, 1–51.
- Blossfeld, H., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D., Prenzel, M. & Wößmann, L. (2009). *Geschlechterdifferenzen im Bildungssystem – Die Bundesländer im Vergleich (Jahresgutachten 2009 des Aktionsrats Bildung)*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bornstein, G. (2003). Intergroup Conflict: Individual, Group, and Collective Interests. *Personality and Social Psychology Review*, 7, 129–145.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bos, W., Hornberg, S., Arnold, K., Faust, G., Fried, L., Lankes, E., Schwippert, K. & Valtin, R. (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (2012a). *IGLU 2011 Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (2012b). *TIMSS 2011 Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Branscombe, N., Ellemers, N., Spears, R. & Doosje, B. (1999). The Context and Content of Social Identity Threat. In N. Ellemers, R. Spears & B. Doosje (Hrsg.), *Social Identity* (S. 35-59). Oxford: Blackwell.
- Branscombe N. & Wann, D. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology*, 24, 641–657.
- Brehm, J. (1966). *A theory of psychological reactance*. New York, NY: Academic Press.
- Brewer, M. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, 17, 475–482.
- Brewer, M. (1979). In-Group Bias in the Minimal Intergroup Situation: A Cognitive-Motivational Analysis. *Psychological Bulletin*, 86, 307–324.
- Brewer, M. & Pickett, C. (2002). The Social Self and Group Identification: Inclusion and Distinctiveness Motives in Interpersonal and Collective Identities. In J. Forgas & K. Williams (Hrsg.), *Sydney Symposium of Social Psychology. The Social Self. Cognitive, Interpersonal and Intergroup Perspectives* (S. 255–272). Hoboken: Taylor and Francis.
- Brown, R., Hinkle, S., Ely, P., Fox-Cardamone, L., Maras, P. & Taylor, L. (1992). Recognizing group diversity: Individualist-collectivist and autonomous-relational social orientations and their implications for intergroup processes. *British Journal of Social Psychology*, 31, 327–342.
- Brown, R. & Josephs, R. (1999). A Burden of Proof: Stereotype Relevance and Gender Differences in Math Performance. *Journal of Personality and Social Psychology*, 76, 246–257.
- Budde, J. (2005). *Männlichkeit und gymnasialer Alltag: Doing gender im heutigen Bildungssystem*. Bielefeld: Transcript.
- Budde, J., Scholand, B. & Faulstich-Wieland, H. (2008). *Geschlechtergerechtigkeit in der Schule: Eine Studie zu Chancen, Blockaden und Perspektiven einer gender-sensiblen Schulkultur*. Weinheim, München: Juventa.
- Budnik, I. & Skale, N. (2007). Integration verhaltensauffälliger Kinder in Grundschulen: Möglichkeiten schulinterner Beratung. *Evangelische Jugendhilfe*, 4, 234–253.
- Buhl, T. (1999). Positive-Negative Asymmetry in Social Discrimination: Meta-Analytical Evidence. *Group Processes & Intergroup Relations*, 2, 51–58.

- Buunk, B., Collins, R., Taylor, S., Van Yperen, N. & Dakof, G. (1990). The Affective Consequences of Social Comparison: Either Direction Has Its Ups and Downs. *Journal of Personality and Social Psychology*, *59*, 1238–1249.
- Buunk, B., Zurriaga, R., Gonzalez-Roma, V. & Subirats, M. (2003). Engaging in upward and downward comparisons as a determinant of relative deprivation at work: A longitudinal study. *Journal of Vocational Behavior*, *62*, 370–388.
- 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., Lombardo, M. & Frigerio, S. (2006). Stereotype threat: the moderating role of Locus of Control beliefs. *European Journal of Social Psychology*, *36*, 183–197.
- 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.
- Carrington, B. (2002). A quintessentially feminine domain? Student teachers' constructions of primary teaching as a career. *Educational Studies*, *28*, 287–303.
- Carrington, B., Tymms, P. & Merrell, C. (2008). Role models, school improvement and the 'gender gap' – do men bring out the best in boys and women the best in girls? *British Educational Research Journal*, *34*, 315–327.
- Causey, D. & Dubow, E. (1992). Development of a Self-Report Coping Measure for Elementary School Children. *Journal of Clinical Child Psychology*, *21*, 47–59.
- Chatard, A., Guimond, S. & Selimbegovic, L. (2007). "How good are you in math?" The effect of gender stereotypes on students' recollection of their school marks. *Journal of Experimental Social Psychology*, *43*, 1017–1024.
- Chatard, A., Selimbegović, L., Konan, P. & Mugny, G. (2008). Performance boosts in the classroom: Stereotype endorsement and prejudice moderate stereotype lift. *Journal of Experimental Social Psychology*, *44*, 1421–1424.
- Cohen, G. & Garcia, J. (2008). Identity, Belonging, and Achievement: A Model, Interventions, Implications. *Current Directions in Psychological Science*, *17*, 365–369.
- Cohen, G. & Garcia, J. (2005). "I Am Us": Negative Stereotypes as Collective Threats. *Journal of Personality and Social Psychology*, *89*, 566–582.
- Cohen, G., Garcia, J., Apfel, N. & Master, A. (2006). Reducing the Racial Achievement Gap: A Social-Psychological Intervention. *Science*, *313*, 1307–1310.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- Collins, R. (1996). For Better or Worse: The Impact of Upward Social Comparison on Self-Evaluations. *Psychological Bulletin*, *119*, 51–69.
- Connell, R. (2000). *Der gemachte Mann: Konstruktion und Krise von Männlichkeiten*. Opladen: Leske und Budrich.
- Cornelißen, W. (2004). *Bildung und Geschlechterordnung in Deutschland.: Einige Anmerkungen zur Debatte um die Benachteiligung von Jungen in der Schule*. München: Deutsches Jugendinstitut e. V.
- Crocker, J. & Luhtanen, R. (1990). Collective Self-Esteem and Ingroup Bias. *Journal of Personality and Social Psychology*, *58*, 60–67.

- Croizet, J., Després, G., Gauzins, M., Huguet, P., Leyens, J. & Méot, A. (2004). Stereotype Threat Undermines Intellectual Performance by Triggering a Disruptive Mental Load. *Personality and Social Psychology Bulletin*, 30, 721–731.
- Cross, S. & Madson, L. (1997). Models of the Self: Self-Construals and Gender. *Psychological Bulletin*, 122, 5–37.
- Crowe, E. & Higgins, E. (1997). Regulatory Focus and Strategic Inclinations: Promotion and Prevention in Decision-Making. *Organizational Behavior and Human Decision Processes*, 69, 117–132.
- Daneman, M. & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Dardenne, B., Dumont, M. & Bollier, T. (2007). Insidious Dangers of Benevolent Sexism: Consequences for Women's Performance. *Journal of Personality and Social Psychology*, 93, 764–779.
- Davies, P., Spencer, S., Quinn, D. & Gerhardstein, R. (2002). Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally. *Personality and Social Psychology Bulletin*, 28, 1615–1628.
- de Lemus, S., Moya, M. & Glick, P. (2010). When Contact Correlates with Prejudice: Adolescents' Romantic Relationship Experience Predicts Greater Benevolent Sexism in Boys and Hostile Sexism in Girls. *Sex Roles*, 63, 214–225.
- Deaux, K. (1985). Sex and Gender. *Annual Review of Psychology*, 36, 49–81.
- Dee, T. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42, 528–554.
- Dee, T. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95, 158–165.
- Devine, P. (1989). Stereotypes and Prejudice: Their Automatic and Controlled Components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Diefenbach, H. & Klein, M. (2002). "Bringing boys back in". Soziale Ungleichheit zwischen den Geschlechtern im Bildungssystem zuungunsten von Jungen am Beispiel der Sekundarschulabschlüsse. *Zeitschrift für Pädagogik*, 48, 938–958.
- Diekmann, A. & Eagly, A. (2000). Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Personality and Social Psychology Bulletin*, 26, 1171–1188.
- Diekmann, A. & Goodfriend, W. (2006). Rolling with the Changes: A Role Congruity Perspective on Gender Norms. *Psychology of Women Quarterly*, 30, 369–383.
- Diener, C. & Dweck, C. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36, 451–462.
- Ditton, H. (2007). *Kompetenzaufbau und Laufbahnen im Schulsystem: Ergebnisse einer Längsschnittuntersuchung an Grundschulen*. Münster: Waxmann.
- Downey, D. & Vogt Yuan, A. (2005). Sex Differences in School Performance during High School: Puzzling Patterns and Possible Explanations. *The Sociological Quarterly*, 46, 299–321.
- Duckworth, A. & Seligman, M. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98, 198–208.
- Dunham, Y., Baron, A. & Carey, S. (2011). Consequences of "Minimal" Group Affiliations in Children. *Child Development*, 82, 793–811.
- Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040–1048.

- Dweck, C. & Elliott, E. (1983). Achievement motivation. In E. Hetherington (Hrsg.), *Socialization, personality, and social development* (S. 643–691). New York, NY: Wiley.
- Dweck, C. & Leggett, E. (1988). A Social-Cognitive Approach to Motivation and Personality. *Psychological Review*, *95*, 256–273.
- Eagly, A. (1987). *Sex differences in social behavior: A social-role interpretation*: Hillsdale, NJ: Lawrence Erlbaum Associates.
- Eagly, A. & Steffen, V. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*, 309–330.
- Eagly, A. & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, *54*, 408–423.
- Eagly, A., Wood, W. & Diekmann, A. (2000). Social Role Theory of Sex Differences and Similarities: A Current Appraisal. In T. Eckes & H. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 123–174). Mahwah, NJ: Lawrence Erlbaum Associates.
- Eccles, J. & Harold, R. (1991). Gender differences in sport involvement: Applying the Eccles' expectancy-value model. *Journal of Applied Sport Psychology*, *3*, 7–35.
- Eccles, J., Jacob, J. & Harold, R. (1990). Gender Role Stereotypes, Expectancy Effects, and Parents' Socialization of Gender Differences. *Journal of Social Issues*, *46*, 183–201.
- Eckert, C. & Imhof, M. (2013). Was Mädchen schadet muss für Jungen noch lange nicht schädlich sein: die Leseleistung von Jungen unter Stereotype Threat. *GENDER – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, *5*, 60–76.
- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker & B. Kortendieck (Hrsg.), *Handbuch Frauen- und Geschlechterforschung* (S. 171–182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Eilders, C., Neidhardt, F. & Pfetsch, B. (1997). Pressekommentare und öffentliche Meinung. In H. Schatz, O. Jarren & B. Knaup (Hrsg.), *Machtkonzentration in der Multimediagesellschaft? Beiträge zu einer Neubestimmung des Verhältnisses von politischer und medialer Macht* (S. 176–187). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ellemers, N. (1993). The Influence of Socio-structural Variables on Identity Management Strategies. *European Review of Social Psychology*, *4*, 27–57.
- Ellemers, N., Wilke, H. & van Knippenberg, A. (1993). Effects of the Legitimacy of Low Group or Individual Status on Individual and Collective Status-Enhancement Strategies. *Journal of Personality and Social Psychology*, *64*, 766–778.
- Ellemers, N. & van Rijswijk, W. (1997). Identity Needs versus Social Opportunities: The Use of Group-Level and Individual-Level Identity Management Strategies. *Social Psychology Quarterly*, *60*, 52–65.
- Elliot, A. (1999). Approach and Avoidance Motivation and Achievement Goals. *Educational Psychologist*, *34*, 169–189.
- Elliot, A. & Church, M. (2003). A Motivational Analysis of Defensive Pessimism and Self-Handicapping. *Journal of Personality*, *71*, 369–396.
- Elliot, A. & Harackiewicz, J. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*, 461–475.
- Elliot, A. & McGregor, H. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519.
- Elliot, A., McGregor, H. & Gable, S. (1999). Achievement Goals, Study Strategies, and Exam Performance: A Mediational Analysis. *Journal of Educational Psychology*, *91*, 549–563.

- Elsbach, K. & Kramer, R. (1996). Members' responses to Organizational Identity Threats: Encountering and Countering the Business Week Rankings. *Administrative Science Quarterly*, 41, 442–476.
- Eschenbeck, H., Kohlmann, C. & Lohaus, A. (2007). Gender Differences in Coping Strategies in Children and Adolescents. *Journal of Individual Differences*, 28, 18–26.
- Evans, L. & Petty, R. (2003). Self-Guide Framing and Persuasion: Responsibly Increasing Message Processing to Ideal Levels. *Personality and Social Psychology Bulletin*, 29, 313–324.
- Faulstich-Wieland, H., Weber, M. & Willems, K. (2004). *Doing Gender im heutigen Schulalltag: Empirische Studien zur sozialen Konstruktion von Geschlecht in schulischen Interaktionen*. Weinheim: Juventa.
- Fend, H. (2001). *Entwicklungspsychologie des Jugendalters: Ein Lehrbuch für pädagogische und psychologische Berufe*. Opladen: Leske und Budrich.
- Ferguson, C. & Kelley, H. (1964). Significant factors in overevaluation of own-group's product. *The Journal of Abnormal and Social Psychology*, 69, 223–228.
- Ferrer-Caja, E. & Weiss, M. (2002). Cross-Validation of a Model of Intrinsic Motivation With Students Enrolled in High School Elective Courses. *The Journal of Experimental Education*, 71, 41–65.
- Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, 7, 117–140.
- Fiske, S., Cuddy, A., Glick, P. & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.
- Fiske, S. & Glick, P. (1995). Ambivalence and Stereotypes Cause Sexual Harassment: A Theory with Implications for Organizational Change. *Journal of Social Issues*, 51, 97–115.
- Fogliati, V. & Bussey, K. (2013). Stereotype Threat Reduces Motivation to Improve: Effects of Stereotype Threat and Feedback on Women's Intentions to Improve Mathematical Ability. *Psychology of Women Quarterly*, 37, 310–324.
- Forbes, C. & Schmader, T. (2010). Retraining attitudes and stereotypes to affect motivation and cognitive capacity under stereotype threat. *Journal of Personality and Social Psychology*, 99, 740–754.
- Frazier, N. & Sadker, M. (1973). *Sexism in school and society*. New York, NY: Harper & Row.
- Frevert, U. (1999). Frauenpolitik als Menschenrechtspolitik. Entwicklung, Stand und Perspektiven der Gleichberechtigung in Deutschland. In F. Hutter & C. Tessmer, *Menschenrechte und Bürgergesellschaft in Deutschland* (S. 51–69). Wiesbaden: Springer.
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., et al. (2009). *PISA-2006-Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Frey, D., Dauenhimer, D., Parge, O. & Haisch, J. (2001). Die Theorie sozialer Vergleichsprozesse. In D. Frey, M. Baldwin & H. Crott (Hrsg.), *Theorien der Sozialpsychologie* (S. 81–122). Bern: Huber.
- Fritz, C., Morris, P. & Richler, J. (2012). Effect Size Estimates: Current Use, Calculations, and Interpretation. *Journal of Experimental Psychology: General*, 141, 2–18.
- Frosh, S., Phoenix, A. & Pattman, R. (2003). The trouble with boys. *The Psychologist*, 16, 84–87.
- Frost, M. (2011). *A Stereotype Threat Intervention that Examines Causal Attributions, Self-Efficacy and Perceived Faculty Support*. Kingston, RI: University of Rhode Island.
- Gabriel, S. & Gardner, W. (1999). Are There "His" and "Hers" Types of Interdependence?: The Implications of Gender Differences in Collective Versus Relational Interdependence for Affect, Behavior, and Cognition. *Journal of Personality and Social Psychology*, 77, 642–655.
- Gärtner, S., Dovidio, J., Anastasio, P., Bachman, B. & Rust, M. (1993). The Common Ingroup Identity Model: Recategorization and the Reduction of Intergroup Bias. *European Review of Social Psychology*, 4, 1–26.

- Gagnon, A. & Bourhis, R. (1996). Discrimination in the Minimal Group Paradigm: Social Identity or Self-Interest? *Personality and Social Psychology Bulletin*, 22, 1289–1301.
- Galambos, N., Almeida, D. & Petersen, A. (1990). Masculinity, Femininity, and Sex Role Attitudes in Early Adolescence: Exploring Gender Intensification. *Child Development*, 6, 1905–1914.
- Gavin, L. & Furman, W. (1989). Age Differences in Adolescents' Perceptions of Their Peer Groups. *Developmental Psychology*, 25, 827–834.
- Gibbons, F. (1986). Social Comparison and Depression: Company's Effect on Misery. *Journal of Personality and Social Psychology*, 51, 140–148.
- Gibbons, F. & Buunk, B. (1999). Individual Differences in Social Comparison: Development of a Scale of Social Comparison Orientation. *Journal of Personality and Social Psychology*, 76, 129–142.
- Gibbons, F. & Gerrard, M. (1991). Downward comparison and coping with threat. In J. Suls & T. Wills (Hrsg.), *Social comparison: Contemporary theory and research* (S. 317–345). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gibbons, F. & Gerrard, M. (1989). Effects of Upward and Downward Social Comparison on Mood States. *Journal of Social and Clinical Psychology*, 8, 14–31
- Gildemeister, R. (2004). Doing Gender: Soziale Praktiken der Doing Gender: Soziale Praktiken der Geschlechterunterscheidung. In R. Becker & B. Kortendiek (Hrsg.), *Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie* (S. 132–140). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gill, R. (2003). Power and the production of subjects: A genealogy of the new man and the new lad. *The Sociological Review*, 51, 34–56.
- Glick, P. & Fiske, S. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, 56, 109–118
- Glick, P. & Fiske, S. (1999). The ambivalence toward men inventory. *Psychology of Women Quarterly*, 23, 519–536.
- Glick, P. & Fiske, S. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491–512.
- Glick, P. & Hilt, L. (2000). From combative children to ambivalent adults: The development of gender prejudice. In T. Eckes & M. Trautner (Hrsg.), *Developmental Social Psychology of Gender* (S. 243–271). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Goff, P., Steele, C. & Davies, P. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology*, 94, 91–107.
- Gómez, Á. (2002). If my group stereotypes others, others stereotype my group...and we know. Concept, research lines and future perspectives of meta-stereotypes. *Revista de Psicología Social*, 17, 253–282.
- Gordijn, E. (2002). Meta-stereotypes and meta-prejudice: Some suggestions for future research. *Revista de Psicología Social*, 17, 283–286.
- Grant, P. (1993). Ethnocentrism in response to a threat to social identity. *Journal of Social Behavior & Personality*, 8, 143–154.
- Greenwald, A., McGhee, D. & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Grundmann, G. & Kramer, R. (2001). Demokratische Reformhoffnungen zwischen schulischen Gestaltungsmöglichkeiten und strukturellen Bedingungen. In J. Böhme & R. Kramer (Hrsg.), *Partizipation in der Schule* (S. 59–92). Opladen: Leske und Budrich.

- Guimond, S., Chatard, A., Martinot, D., Crisp, R. & Redersdorff, S. (2006). Social Comparison, Self-Stereotyping, and Gender Differences in Self-Construals. *Journal of Personality and Social Psychology*, *90*, 221–242.
- Gummerum, M., Keller, M., Takezawa, M. & Mata, J. (2008). To give or not to give: Children's and adolescents' sharing and moral negotiations in economic decision situations. *Child Development*, *79*, 562–576.
- Gupta, V. & Bhawe, N. (2007). The influence of proactive personality and stereotype threat on women's entrepreneurial intentions. *Journal of Leadership & Organizational Studies*, *13*, 73–85.
- Gupta, V., Turban, D. & Bhawe, N. (2008). The effect of gender stereotype activation on entrepreneurial intentions. *Journal of Applied Psychology*, *93*, 1053–1061.
- Hadjar, A. & Lupatsch, J. (2011). Geschlechterunterschiede im Schulerfolg: Spielt die Lehrperson eine Rolle? *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *31*, 79–94.
- Hannover, B. & Greve, W. (2012). Selbst und Persönlichkeit. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (S. 543–561). Weinheim: Beltz.
- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, *25*, 89–103.
- Hannover, B. & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why German high school students do not like math and science. *Learning and Instruction*, *14*, 51–67.
- Hannover, B. & Kessels, U. (2002). Challenge the science-stereotype. Der Einfluss von Technik-Freizeitkursen auf das Naturwissenschaften-Stereotyp von Schülerinnen und Schülern. In M. Prenzel & J. Doll (Hrsg.), *Zeitschrift für Pädagogik, Beiheft 45. Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 341–358). Weinheim: Beltz.
- Harackiewicz, J. & Elliot, A. (1993). Achievement goals and intrinsic motivation. *Journal of Personality and Social Psychology*, *65*, 904–915.
- Harter, S. (2003). The development of self-representations during childhood and adolescence. In M. Leary & J. Tangney (Hrsg.), *Handbook of Self and Identity* (S. 610–642). New York, NY: Guilford Press.
- Harter, S. (1990). Identity and self development. In S. Feldman & G. Elliott (Hrsg.), *At the threshold. The developing adolescent* (S. 352–387). Cambridge, MA: Harvard University Press.
- Harter, S., Bresnick, S., Bouchey, H. & Whitesell, N. (1997). The development of multiple role-related selves during adolescence. *Development and Psychopathology*, *9*, 835–853.
- Hartley, B. & Sutton, R. (2013). A Stereotype Threat Account of Boys' Academic Underachievement. *Child Development*, *84*, 1716–1733.
- Haslam, S., Ellemers, N., Reicher, S., Reynolds, K. & Schmitt, M. (2010). Social identity today: The impact of its defining ideas. In T. Postmes & N. Branscombe (Hrsg.), *Rediscovering social identity: Core sources*. New York, NY: Psychology Press.
- Haslam, S. & Turner, J. (1992). Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation. *European Journal of Social Psychology*, *22*, 251–277.
- Haslam, S., Turner, J., Oakes, P., McGarty, C. & Hayes, B. (1992). Context-dependent variation in social stereotyping 1: The effects of intergroup relations as mediated by social change and frame of reference. *European Journal of Social Psychology*, *22*, 3–20.
- Heisig, J. (2011). Who Does More Housework: Rich or Poor?: A Comparison of 33 Countries. *American Sociological Review*, *76*, 74–99.

- Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 93–111.
- Heyder, A. & Kessels, U. (2013). Is School Feminine? Implicit Gender Stereotyping of School as a Predictor of Academic Achievement. *Sex Roles*, 69, 605–617.
- Heyder, A. & Kessels, U. (2012, September). *Harter Junge = fauler Junge? Eine experimentelle Studie zur Wirkung von Männlichkeitsinszenierungen in der Schule auf Lehrkräfte*. Vortrag auf dem 47. Kongress der Deutschen Gesellschaft für Psychologie, Bielefeld.
- Hidi, S. & Renninger, K. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, 41, 111–127.
- Hirstein, M., Freund, N. & Hausmann, M. (2012). Gender Stereotyping Enhances Verbal Fluency Performance in Men (and Women). *Zeitschrift für Psychologie*, 220, 70–77.
- Hogg, M. & Mullin, B. (1999). Joining groups to reduce uncertainty: Subjective uncertainty reduction and group identification. In D. Abrams & M. Hogg (Hrsg.), *Social identity and social cognition* (S. 249–279). Malden, MA: Blackwell.
- Hogg, M. & Terry, D. (2001). *Social Identity Processes in organizational Contexts*. Philadelphia: Psychology Press.
- Hogg, M. & Williams, K. (2000). From I to We: Social identity and the collective self. *Group Dynamics: Theory, Research, and Practice*, 4, 81–97.
- Holst, E. & Wiemer, A. (2010). Women still greatly underrepresented on the top boards of large companies. *German Institute for Economic Research Weekly Report*, 6, 45–53.
- Holz-Ebeling, F., Grätz-Tümmers, J. & Schwarz, C. (2000). Jungen als "Nutznießer" der Koedukation?: Eine empirische Studie zur Bedeutung der Koedukation für Jungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32, 94–107.
- Huang, C. (2012). Discriminant and Criterion-Related Validity of Achievement Goals in Predicting Academic Achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104, 48–73.
- Huguet, P., Dumas, F., Monteil, J. & Genestoux, N. (2001). Social comparison choices in the classroom: further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology*, 31, 557–578.
- Huguet, P. & Régner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of Experimental Social Psychology*, 45, 1024–1027.
- Hyde, J. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592.
- Hyde, J. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology*, 20, 722–736.
- Hyde, J., Fennema, E., Ryan, M., Frost, L. & Hopp, C. (1990). Gender Comparisons of Mathematics Attitudes and Affect. A Meta-Analysis. *Psychology of Women Quarterly*, 14, 299–324.
- Institut zur Qualitätsentwicklung im Bildungswesen (IQB) (2008). *Vergleichsarbeiten 2008, 8. Jahrgangsstufe: Testhefte I & II, Deutsch & Mathematik*. Berlin: IQB.
- Inzlicht, M. & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11, 365–371.
- Inzlicht, M., Tullett, M. & Gutsell, J. (2011). Stereotype Threat Spillover: The Short- and Long-Term Effects of Coping with Threats to Social Identity. In M. Inzlicht & T. Schmader (Hrsg.), *Stereotype Threat: Theory, Process, and Application* (S. 107–123). Oxford, NY: University Press.
- Issmer, C., Stellmacher, J. & Gollwitzer, M. (2013). When disadvantaged adolescents strike out: the impact of negative metastereotypes on delinquency. *Journal of Criminal Psychology*, 3, 4–18.

- Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy, *Gender and Education*, 14, 37–51.
- Jackson, J. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *The Psychological Record*, 43, 395–413.
- Jackson, J. & Smith, E. (1999). Conceptualizing Social Identity: A New Framework and Evidence for the Impact of Different Dimensions. *Personality and Social Psychology Bulletin*, 25, 120–135.
- Jackson, L., Sullivan, L., Harnish, R. & Hodge, C. (1996). Achieving Positive Social Identity: Social Mobility, Social Creativity, and Permeability of Group Boundaries. *Journal of Personality and Social Psychology*, 70, 241–254.
- Jetten, J., Spears, R. & Manstead, A. (1999). Group Distinctiveness and Intergroup Discrimination. In N. Ellemers, R. Spears & B. Doosje (Hrsg.), *Social Identity* (S. 107–126). Oxford: Blackwell.
- Jetten, J., Spears, R. & Manstead, A. (1997). Distinctiveness threat and prototypicality: combined effects on intergroup discrimination and collective self-esteem. *European Journal of Social Psychology*, 27, 635–657.
- Jetten, J., Spears, R. & Manstead, A. (1996). Intergroup Norms and Intergroup Discrimination: Distinctive Self-Categorization and Social Identity Effects. *Journal of Personality and Social Psychology*, 71, 1222–1233.
- Johns, M., Inzlicht, M. & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, 137, 691–705.
- Johns, M., Schmader, T. & Martens, A. (2005). Knowing Is Half the Battle: Teaching Stereotype Threat as a Means of Improving Women's Math Performance. *Psychological Science*, 16, 175–179.
- Johnson, H., Barnard-Brak, L., Saxon, T. & Johnson, M. (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.
- Jones, D. (2001). Social Comparison and Body Image: Attractiveness Comparisons to Models and Peers Among Adolescent Girls and Boys. *Sex Roles*, 45, 645–664.
- Jones, S. & Myhill, D. (2004). Troublesome boys' and 'compliant girls': gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, 25, 547–561.
- Kahnemann, D. & Miller, D. (1986). Norm Theory: Comparing Reality to Its Alternatives. *Psychological Review*, 93, 136–153.
- Kaufmann, A. (2007). Merkmale und Einstellungen von Schülern. In H. Ditton (Hrsg.), *Kompetenzaufbau und Laufbahnen im Schulsystem* (S. 117–144). Münster: Waxmann.
- Keil, S. (2013). Soziale Partizipation. In J. van Deth & M. Tausendpfund (Hrsg.), *Politik im Kontext: Ist alle Politik lokale Politik?* (S. 157–189). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Keller, J. (2007). When Negative Stereotypic Expectancies Turn into Challenge or Threat: The Moderating Role of Regulatory Focus. *Swiss Journal of Psychology*, 66, 163–168.
- Keller, J. (2002). Blatant Stereotype Threat and Women's Math Performance: Self-Handicapping as a Strategic Means to Cope with Obtrusive Negative Performance Expectations. *Sex Roles*, 47, 193–198.
- Kempf, W. (1974). Dynamische Modelle zur Messung sozialer Verhaltenspositionen. In W. Kempf (Hrsg.), *Probabilistische Modelle in der Sozialpsychologie* (S. 13–55). Bern: Huber.
- Kernis, M. & Goldman, B. (2003). Stability and Variability in Self-Concept and Self-Esteem. In M. Leary & J. Tangney (Hrsg.), *Handbook of Self and Identity* (S. 106–127). New York: Guilford Press.
- Kessels, U. (2012). Selbstkonzept: Geschlechtsunterschiede und Interventionsmöglichkeiten. In H. Stöger, A. Ziegler & M. Heilemann (Hrsg.), *Mädchen und Frauen in MINT. Bedingungen von Geschlechtsunterschieden und Interventionsmöglichkeiten* (S. 163–192). Münster: LIT Verlag.

- Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education, 20*, 309–323.
- Kessels, U. (2002). *Undoing gender in der Schule: Eine empirische Studie über Koedukation und Geschlechtsidentität im Physikunterricht. Materialien*. Weinheim: Juventa.
- Kessels, U. & Hannover, B. (2006). Zum Einfluss des Image von mathematisch-naturwissenschaftlichen Schulfächern auf die schulische Interessenentwicklung. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 350–369). Münster: Waxmann.
- Kessels, U. & Hannover, B. (2004). Entwicklung schulischer Interessen als Identitätsregulation. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 398–412). Münster: Waxmann.
- Kessels, U. & Steinmayr, R. (2013). Macho-man in school: Toward the role of gender role self-concepts and help seeking in school performance. *Learning and Individual Differences, 23*, 234–240.
- Kiefer, A. (2006). *The effects of stereotypes on performance attributions: How gender-math stereotypes lead women to internalize failure*. Ann Arbor, MI: University of Michigan.
- Kiefer, A. & Sekaquaptewa, D. (2007). Implicit Stereotypes, Gender Identification, and Math-Related Outcomes. A Prospective Study of Female College Students. *Psychological Science, 18*, 13–18.
- Kiefer, A. & Shih, M. (2006). Gender differences in persistence and attributions in stereotype relevant contexts. *Sex Roles, 54*, 859–868.
- Kilmartin, C., Smith, T., Green, A., Heinzen, H., Kuchler, M. & Kolar, D. (2008). A Real Time Social Norms Intervention to Reduce Male Sexism. *Sex Roles, 59*, 264–273.
- Kite, M., Deaux, K. & Haines, E. (2008). Gender Stereotypes. In F. Denmark & M. Paludi (Hrsg.), *Psychology of Women: A Handbook of Issues and Theories* (S. 205–236). Westport, CT: Praeger.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M. et al. (2010). *PISA 2009: Bilanz nach einem Jahrzehnt*. Münster: Waxmann.
- Klingler, W. (2008). Jugendliche und ihre Mediennutzung 1998 bis 2008. *Media Perspektiven, 12*, 625–634.
- Knigge, M. & Hannover, B. (2011). Collective school-type identity: predicting students' motivation beyond academic self-concept. *International journal of psychology, 46*, 191–205.
- Koch, S., Müller, S. & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education, 51*, 1795–1803.
- Koenig, A. & Eagly, A. (2005). Stereotype Threat in Men on a Test of Social Sensitivity. *Sex Roles, 52*, 489–496.
- Koenig, A., Eagly, A., Mitchell, A. & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin, 137*, 616–642.
- Kölbl, C., Tiedemann, J. & Billmann-Mahecha, E. (2006). Die Bedeutung der Lesekompetenz für Sachfächer. *Psychologie in Erziehung und Unterricht, 53*, 201–212.
- Konsortium Bildungsberichterstattung (2006). *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.
- Krahé, B., Busching, R. & Möller, I. (2012). Media violence use and aggression among German adolescents: Associations and trajectories of change in a three-wave longitudinal study. *Psychology of Popular Media Culture, 1*, 152–166.
- Kray, L., Reb, J., Galinsky, A. & Thompson, L. (2004). Stereotype Reactance at the Bargaining Table: The Effect of Stereotype Activation and Power on Claiming and Creating Value. *Personality and Social Psychology Bulletin, 30*, 399–411.

- Kray, L., Thompson, L. & Galinsky, A. (2001). Battle of the Sexes: Gender Stereotype Confirmation and Reactance in Negotiations. *Journal of Personality and Social Psychology*, 80, 942–958.
- Krebs, A. (2008). *Jungen erleben Schule: Personzentrierte Jungenforschung: Methodik, Ergebnisse und Perspektiven schulische Jungenarbeit*. München: M-Press.
- Krebs, A. (2002). *Sichtweisen und Einstellungen heranwachsender Jungen: Ergebnisse einer Befragung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Kruskal, W. & Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47, 583–621.
- Kuhl, J. (1987). Action control: The maintenance of motivational states. In F. Halisch & J. Kuhl, *Motivation, intention, and volition* (S. 279–291). Berlin: Springer.
- Kuhl, P. & Hannover, B. (2012). Differenzielle Benotungen von Mädchen und Jungen. Der Einfluss der von der Lehrkraft eingeschätzten Kompetenz zum selbstgesteuerten Lernen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 153–162.
- Kunter, M. & Stanat, P. (2002). Soziale Kompetenz von Schülerinnen und Schülern. *Zeitschrift für Erziehungswissenschaft*, 5, 49–71.
- Labouvie-Vief, G. (2003). Dynamic integration: affect, cognition, and the self in adulthood. *Current Directions in Psychological Science*, 12, 201–206.
- Lahelma, E. (2000). Lack of male teachers: A problem for students or teachers? *Pedagogy, Culture and Society*, 8, 173–186.
- Larson, R. (2000). Toward a Psychology of Positive Youth Development. *American Psychologist*, 55, 170–183.
- Latrofa, M., Vaes, J. & Cadinu, M. (2012). Self-Stereotyping: The Central Role of an Ingroup Threatening Identity. *The Journal of Social Psychology*, 152, 92–111.
- Latrofa, M., Vaes, J., Cadinu, M. & Carnaghi, A. (2010). The Cognitive Representation of Self-Stereotyping. *Personality and Social Psychology Bulletin*, 36, 911–922.
- Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!? How the discourse on "failing boys" impacts performances and motivational goal orientation in German school students. *Social Psychology*, 45, 112–126.
- Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C. (1981). *Das State-Trait-Angstinventar: Theoretische Grundlagen und Handanweisung*. Weinheim: Beltz Test GmbH.
- Leggett, E. (1985, März). *Children's entity and incremental theories of intelligence: Relationships to achievement behavior*. Paper presented at the annual meeting of the Eastern Psychological Association, Boston.
- Lemyre, L. & Smith, P. (1985). Intergroup discrimination and self-esteem in the minimal group paradigm. *Journal of Personality and Social Psychology*, 49, 660–670.
- Lesko, A. & Corpus, J. (2006). Discounting the Difficult: How High Math-Identified Women Respond to Stereotype Threat. *Sex Roles*, 54, 113–125.
- Leuze, K. & Strauß, S. (2009). Lohnungleichheiten zwischen Akademikerinnen und Akademikern: Der Einfluss von fachlicher Spezialisierung, frauendominierten Fächern und beruflicher Segregation. *Zeitschrift für Soziologie*, 38, 262–281.
- Levy, B. (1996). Improving memory in old age by implicit self-stereotyping. *Journal of Personality and Social Psychology*, 71, 1092–1107.
- Leyens, J., Désert, M., Croizet, J. & Darcis, C. (2000). Stereotype Threat: Are Lower Status and History of Stigmatization Preconditions of Stereotype Threat? *Personality and Social Psychology Bulletin*, 26, 1189–1199.

- Lindeman, M. (1997). Ingroup bias, self-enhancement and group identification. *European Journal of Social Psychology, 27*, 337–355.
- Lindeman, M., Harakka, T. & Keltikangas-Järvinen, L. (1997). Age and Gender Differences in Adolescents' Reactions to Conflict Situations: Aggression, Prosociality, and Withdrawal. *Journal of Youth and Adolescence, 26*, 339–351.
- Lockwood, P. (2002). Could It Happen to You? Predicting the Impact of Downward Comparisons on the Self. *Journal of Personality and Social Psychology, 82*, 343–358.
- Lörcher, G. & Maier, P. (1999). *Was erreichen Schüler und Lehrer im Fach Mathematik*. Unveröffentlichtes Manuskript: Pädagogische Hochschule Freiburg.
- Luddy, J. & Thompson, E. (1997). Masculinities and Violence: A Father-Son Comparison of Gender Traditionality and Perceptions of Heterosexual Rape. *Journal of Family Psychology, 11*, 462–477.
- Madera, J., Hebl, M. & Martin, R. (2009). Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology, 94*, 1591–1599.
- Major, B. & O'Brien, L. (2005). The Social Psychology of Stigma. *Annual Review of Psychology, 56*, 393–421.
- Major, B., Quinton, W., McCoy, S. & Schmader, T. (2000). Reducing Prejudice: The Target's Perspective. In S. Oskamp (Hrsg.), *Reducing prejudice and discrimination* (S. 211–238). New York, NY: Psychology Press.
- Major, B., Spencer, S., Schmader, T., Wolfe, C. & Crocker, J. (1998). Coping with Negative Stereotypes about Intellectual Performance: The Role of Psychological Disengagement. *Personality and Social Psychology Bulletin, 24*, 34–50.
- Marsh, H. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*, 77–172.
- Marsh, H. & Ayotte, V. (2003). Do Multiple Dimensions of Self-Concept Become More Differentiated With Age? The Differential Distinctiveness Hypothesis. *Journal of Educational Psychology, 95*, 687–706.
- Marsh, H., Martin, A. & Cheng, J. (2008). A multilevel perspective on gender in classroom motivation and climate: Potential benefits of male teachers for boys? *Journal of Educational Psychology, 100*, 78–95.
- Martin, A. (2004). School motivation of boys and girls: Difference of degree, difference of kind, or both? *Australian Journal of Psychology, 56*, 133–146.
- Martin, A. & Marsh, H. (2005). Motivating Boys and Motivating Girls: Does Teacher Gender Really Make a Difference? *Australian Journal of Education, 49*, 320–334.
- Martin, C., Ruble, D. & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*, 903–933.
- Martiny, S. & 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. 153–178). Berlin, Münster: LIT-Verl.
- Martiny, S., Roth, J. & Deaux, K. (2011). The role of domain identification and arousal in stereotype threat. Manuscript in Vorbereitung.
- Martiny, S., Roth, J., Jelenec, P., Steffens, M. & Croizet, J. (2012). When a new group identity does harm on the spot: Stereotype threat in newly created groups. *European Journal of Social Psychology, 42*, 65–71.
- McGlone, M. & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486–493.
- McIntyre, R., Paulson, R., Taylor, C., Morin, A. & Lord, C. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology, 41*, 301–311.

- McNatt, D. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology, 85*, 314–322.
- Merton, R. (1948). The Self-Fulfilling Prophecy. *The Antioch Review, 8*, 193–210.
- Michalek, R. (2009): Gruppendiskussionen mit Grundschulern. In J. Budde & I. Mammes (Hrsg.), *Jungenforschung empirisch – Zwischen Schule, männlichem Habitus und Peerkultur* (S. 47–71). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Midgley, C., Maehr, M., Hruda, L., Anderman, E., Anderman, L., Freeman, K. et al. (2000). *Manual for the Patterns of Adaptive Learning Scales*. Ann Arbor, MI: University of Michigan.
- Mlicki, P. & Ellemers, N. (1996). Being different or being better? National stereotypes and identifications of Polish and Dutch students. *European Journal of Social Psychology, 26*, 97–114.
- Molden, D. & Finkel, E. (2010). Motivations for promotion and prevention and the role of trust and commitment in interpersonal forgiveness. *Journal of Experimental Social Psychology, 46*, 255–268.
- Möller, J. & Schiefele, U. (2007). Editorial zum Themenschwerpunkt „Lesekompetenz, Lesemotivation und Lesestrategien“. *Psychologie in Erziehung und Unterricht, 4*, 257–258.
- Mullen, B., Brown, R. & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103–122.
- Mullis, I., Martin, M., Kennedy, A. & Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*. Boston: IEA.
- Mummendey, A. (1985). Verhalten zwischen sozialen Gruppen: Die Theorie der sozialen Identität. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie, Band II: Gruppen- und Lerntheorien* (S. 185–216). Bern: Hans Buber.
- Mummendey, A., Otten, S., Berger, U. & Kessler, T. (2000). Positive-Negative Asymmetry in Social Discrimination: Valence of Evaluation and Salience of Categorization. *Personality and Social Psychology Bulletin, 26*, 1258–1270.
- Mummendey, A., Simon, B., Dietze, C., Grünert, M., Haeger, G., Kessler, S. et al. (1992). Categorization Is Not Enough: Intergroup Discrimination in Negative Outcome Allocation. *Journal of Experimental Social Psychology, 28*, 125–144.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*, 472–489.
- Mutz, M. (2009). Sportbegeisterte Jungen, sportabstinente Mädchen?: Eine quantitative Analyse der Sportvereinszugehörigkeit von Jungen und Mädchen mit ausländischer Herkunft. *Sport und Gesellschaft, 6*, 95–121.
- Neugebauer, M. (2011). Werden Jungen von Lehrerinnen bei den Übergangsempfehlungen für das Gymnasium benachteiligt? Eine Analyse auf Basis der IGLU-Daten. In A. Hadjar (Hrsg.), *Geschlechtsspezifische Bildungsungleichheiten* (S. 235–260). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Neuvill, E. & Croizet, J. (2007). Can salience of gender identity impair math performance among 7–8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education, 22*, 307–316.
- Nguyen, H. & Ryan, A. (2008). Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence. *Journal of Applied Psychology, 93*, 1314–1334.
- Nicholls, J. (1975). Causal Attributions and Other Achievement-Related Cognitions: Effects of Task Outcome, Attainment Value, and Sex. *Journal of Personality and Social Psychology, 31*, 379–389.
- Nosek, B., Banaji, M. & Greenwald, A. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*, 44–59.
- Oakes, P., Haslam, S. & Turner, J. (1994). *Stereotyping and social reality*. Oxford: Blackwell.

- Oakes, P. & Turner, J. (1980). Social categorization and intergroup behaviour: Does minimal intergroup discrimination make social identity more positive? *European Journal of Social Psychology*, 10, 295–301.
- O'Brien, L. & Crandall, C. (2003). Stereotype Threat and Arousal: Effects on Women's Math Performance. *Personality and Social Psychology Bulletin*, 29, 782–789.
- Organisation for Economic Co-operation and Development (OECD) (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD) (2011). *PISA 2009 Ergebnisse: Potenzielle nutzen und Chancengerechtigkeit sichern: Sozialer Hintergrund und Schülerleistungen (Band II)*. Paris: OECD Publishing.
- Osborne, J. & Jones, B. (2011). Identification with Academics and Motivation to Achieve in School: How the Structure of the Self Influences Academic Outcomes. *Educational Psychology Review*, 23, 131–158.
- Oswald, D. & Harvey, R. (2000/2001). Hostile Environments, Stereotype Threat, and Math Performance among Undergraduate Women. *Current Psychology*, 19, 338–356.
- Otis A. & Lennon R. (1997). *Otis-Lennon School Ability Test (OLSAT) technical manual*. San Antonio, TX: Harcourt Brace Educational Measurement.
- Papastergiou, M. (2008). Are Computer Science and Information Technology still masculine fields? High school students' perceptions and career choices. *Computers & Education*, 51, 594–608.
- Parker, L. & Rennie, L. (2002). Teachers' implementation of gender-inclusive instructional strategies in single-sex and mixed-sex science classrooms. *International Journal of Science Education*, 24, 881–897.
- Paulus, P. (2003, November). *Psychische Gesundheit – auch ein Problem von Schulen? Dokumentation des Kongresses Gute und gesunde Schule* (S. 76–99), Düsseldorf.
- Picho, K., Rodriguez, A. & Finnie, L. (2013). Exploring the Moderating Role of Context on the Mathematics Performance of Females under Stereotype Threat: A Meta-Analysis. *The Journal of Social Psychology*, 153, 299–333.
- Pintrich, P. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544–555.
- PISA-Konsortium Deutschland (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- Plante, I., Théorêt, M. & Favreau, O. (2009). Student gender stereotypes: contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology*, 29, 385–405.
- Platow, M., Harley, K., Hunter, J., Hanning, P., Shave, R. & O'Connell, A. (1997). Interpreting ingroup-favouring allocations in the minimal group paradigm. *British Journal of Social Psychology*, 16, 107–117.
- Prentice, D. & Carranza, E. (2002). What Women and Men Should Be, Shouldn't Be, Are Allowed to Be, and Don't Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly*, 26, 269–281.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32, 163–173.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Quenzel, G. & Hurrelmann, K. (2010). Geschlecht und Schulerfolg: Ein soziales Stratifikationsmuster kehrt sich um. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 61–91.
- Raab, G., Unger, A. & Unger, F. (2010). *Marktpsychologie: Grundlagen und Anwendung*. Heidelberg: Springer.

- Rabbie, J. & Wilkens, G. (1971). Intergroup competition and its effect on intragroup and intergroup relations. *European Journal of Social Psychology*, 1, 215–234.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Kopenhagen: Danish Institute for Educational Research.
- Rau, M. (2013). *Geschlechtsbezogene Bildungsdisparitäten: Die Bedeutung der Zuschreibung gendertypisiert-er Merkmale und des ambivalenten Sexismus bei Jugendlichen für ihren Bildungserfolg*. Hamburg: Kovač.
- Rau, M. & Hannover, B. (2008, Juli). *How students' sexist beliefs affect educational success*. Vortrag auf dem XXIX International Congress of Psychology (ICP), Berlin.
- Rawsthorne, L. & Elliot, A. (1999). Achievement Goals and Intrinsic Motivation: A Meta-Analytic Review. *Personality and Social Psychology Review*, 3, 326–344.
- Regner, I., Smeding, A., Gimming, D., Thinus-Blanc, C., Monteil, J. & Huguet, P. (2010). Individual Differences in Working Memory Moderate Stereotype-Threat Effects. *Psychological Science*, 21, 1646–1648.
- Rentzsch, K., Schütz, A. & Schröder-Abé, M. (2011). Being Labeled Nerd: Factors That Influence the Social Acceptance of High-Achieving Students. *The Journal of Experimental Education*, 79, 143–168.
- Rheinberg, F. (2006). *Motivation*. Stuttgart: Kohlhammer.
- Ricciardelli, L., McCabe, M. & Ridge, D. (2006). The Construction of the Adolescent Male Body through Sport. *Journal of Health Psychology*, 11, 577–587.
- Richman, L. & Leary, M. (2009). Reactions to Discrimination, Stigmatization, Ostracism, and Other Forms of Interpersonal Rejection: A Multimotive Model. *Psychological Review*, 116, 365–383.
- Roccas, S. & Schwartz, S. (1993). Effects of intergroup similarity on intergroup relations. *European Journal of Social Psychology*, 23, 581–595.
- Rost, J. (2003). *Lehrbuch Testtheorie – Testkonstruktion*. Göttingen: Huber.
- Rubin, M. & Hewstone, M. (1998). Social Identity Theory's Self-Esteem Hypothesis: A Review and Some Suggestions for Clarification. *Personality and Social Psychology Review*, 2, 40–62.
- Rudman, L. & Glick, P. (1999). Feminized Management and Backlash toward Agentic Women: The Hidden Costs to Women of a Kinder, Gender Image of Middle Managers. *Journal of Personality and Social Psychology*, 77, 1004–1010.
- Rudman, L. & Goodwin, S. (2004). Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More Than Men Like Men? *Journal of Personality and Social Psychology*, 87, 494–509.
- Rudolph, U. (2003). *Motivationspsychologie*. Weinheim: Beltz.
- Ryan, K. & Ryan, A. (2005). Psychological Processes Underlying Stereotype Threat and Standardized Math Test Performance. *Educational Psychologist*, 40, 53–63.
- Schaufler, B. (2002). *"Schöne Frauen, starke Männer": Zur Konstruktion von Leib, Körper und Geschlecht*. Opladen: Leske und Budrich.
- Schiefele, U. (2008). Lernmotivation und Interesse. In M. Hasselhorn (Hrsg.), *Handbuch der Psychologie*, Bd. 10 (S. 38–49). Göttingen: Hogrefe.
- Schimmel, J., Arndt, J., Banko, K. & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition*, 22, 75–99.
- Schlack, R., Hölling, H., Kurth, B. & Huss, M. (2007). Die Prävalenz der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) bei Kindern und Jugendlichen in Deutschland. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 50, 827–835.
- 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*, 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*, *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.
- Schmader, T., Major, B. & Gramzow, R. (2001). Coping With Ethnic Stereotypes in the Academic Domain: Perceived Injustice and Psychological Disengagement. *Journal of Social Issues*, *57*, 93–111.
- Schmalz, D. & Kerstetter, D. (2006). Girlie girls and manly men: children's stigma consciousness of gender in sports and physical activities. *Journal of Leisure Research*, *38*, 536–557.
- Schöps, K., Walter, O., Zimmer, K. & Prenzel, M. (2006). Disparitäten zwischen Jungen und Mädchen in der mathematischen Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 209–224). Münster: Waxmann.
- Schreiter, J. & Bürger, T. (2006). *Praxisbaustein: Klassen- und Schulspercherqualifizierung als zentrales Element einer demokratischen Schulstruktur*. Berlin: Unveröffentlichtes Manuskript.
- Schug, M. & Birkey, C. (1985). The development of children's economic reasoning. *Theory and Research in Social Education*, *13*, 31–42.
- Segeritz, M., Stanat, P. & Walter, O. (2010). Muster des schulischen Erfolgs von Mädchen und Jungen mit Migrationshintergrund. In C. Allemann-Ghionda, P. Stanat, K. Göbel & C. Röhner (Hrsg.), *Zeitschrift für Pädagogik: 55. Beiheft. Migration, Identität, Sprache und Bildungserfolg* (S. 165–186). Weinheim, Basel: Beltz.
- Seibt, B. & Förster, J. (2004). Stereotype Threat and Performance: How Self-Stereotypes Influence Processing by Inducing Regulatory Foci. *Journal of Personality and Social Psychology*, *87*, 38–56.
- Seiffge-Krenke, I. (1994). *Gesundheitspsychologie des Jugendalters*. Göttingen: Hogrefe.
- Senatsverwaltung für Bildung und Forschung (2013). *Fortbildung für das pädagogische Personal*. Berlin: SenBJW.
- Shavelson, R., Hubner, J. & Stanton, G. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research*, *46*, 407–441.
- Sherif, M. (1967). *Group conflict and co-operation: Their social psychology*. London: Routledge.
- Sherif, M., Harvey, O., White, B., Hood, W. & Sherif, C. (1961). *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. Norman, OK: University Book Exchange.
- Shih, M., Ambady, N., Richeson, J., Fujita, K. & Gray, H. (2002). Stereotype Performance Boosts: The Impact of Self-Relevance and the Manner of Stereotype Activation. *Journal of Personality and Social Psychology*, *83*, 638–647.
- Shih, M., Pittinsky, T., Ho, G. (2012). Stereotype Boost: Positive Outcomes from the Activation of Positive Stereotypes. In M. Inzlicht & T. Schmader (Hrsg.), *Stereotype Threat: Theory, Process, and Application* (S. 141–158). New York, NY: Oxford University Press.
- Sibley, C., Wilson, M. & Duckitt, J. (2007). Antecedents of Men's Hostile and Benevolent Sexism: The Dual Roles of Social Dominance Orientation and Right-Wing Authoritarianism. *Personality and Social Psychology Bulletin*, *33*, 160–172.
- Simon, B., Pantaleo, G. & Mummendey, A. (1995). Unique individual or interchangeable group member? The accentuation of intragroup differences versus similarities as an indicator of the individual self versus the collective self. *Journal of Personality and Social Psychology*, *69*, 106–119.
- Simon, B. & Trötschel, R. (2007). Das Selbst und die soziale Identität. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 147–186). Heidelberg: Springer.

- Skelton, C. (2002) The 'feminisation of schooling' or 're-masculinising' primary education, *International Studies in Sociology of Education*, 12, 79–98.
- Skelton, C. & Francis, B. (2011). Successful Boys and Literacy: Are "Literate Boys" Challenging or Repackaging Hegemonic Masculinity? *Curriculum Inquiry*, 41, 456–479.
- Smith, J. (2006). The Interplay among Stereotypes, Performance-Avoidance Goals, and Women's Math Performance Expectations. *Sex Roles*, 54, 287–296.
- Smith, J. (2004). Understanding the Process of Stereotype Threat: A Review of Mediational Variables and New Performance Goal Directions. *Educational Psychology Review*, 16, 177–206.
- Smith, J. & Lewis, K. (2009). Men's Interpersonal (mis)Perception: Fitting in with Gender Norms Following Social Rejection. *Sex Roles*, 61, 252–264.
- Smith, J., Sansone, C. & White, P. (2007). The Stereotyped Task Engagement Process: The Role of Interest and Achievement Motivation. *Journal of Educational Psychology*, 99, 99–114.
- Smith, J. & White, P. (2001). Development of the Domain Identification Measure: A Tool for Investigating Stereotype Threat Effects. *Educational and Psychological Measurement*, 61, 1040–1057.
- Smith, R. (2000). Assimilative and Contrastive Emotional reactions to Upward and Downward Social Comparisons. In J. Suls & L. Wheeler (Hrsg.), *Handbook of Social Comparison. Theory and Research* (S. 173–200). Boston, MA: Springer.
- Spears, R., Doosje, B. & Ellemers, N. (1997). Self-Stereotyping in the Face of Threats to Group Status and Distinctiveness: The Role of Group Identification. *Personality and Social Psychology Bulletin*, 23, 538–553.
- Spencer, S., Steele, C. & Quinn, D. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35, 4–28.
- Spiel, C., Wagner, P. & Fellner, G. (2002). Wie lange arbeiten Kinder zu Hause für die Schule? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 34, 125–135.
- Spielberger, C., Gorsuch, R. & Lushene, R. (1970). *Manual for the state-trait anxiety inventory (STAI)*. Palo Alto, CA: Consulting Psychologists Press.
- Spielman, D. (2000). Young Children, Minimal Groups, and Dichotomous Categorization. *Personality and Social Psychology Bulletin*, 26, 1433–1441.
- Spinath, B. (2009). Zielorientierungen. In V. Brandstätter (Hrsg.), *Handbuch der Allgemeinen Psychologie. Motivation und Emotion* (S. 64–71). Göttingen: Hogrefe.
- Spinath, B., Stiensmeier-Pelster, J., Schöne, C. & Dickhäuser, O. (2002). *SELLMO: Skalen zur Erfassung der Lern- und Leistungsmotivation*. Göttingen: Hogrefe.
- Stamm, M. (2007). Begabung, Leistung und Geschlecht: neue Dimensionen im Lichte eines alten erziehungswissenschaftlichen Diskurses. *Review of Education*, 53, 417–437.
- Stanat, P. & Kunter, M. (2003). Kompetenzerwerb, Bildungsbeteiligung und Schullaufbahn von Mädchen und Jungen im Ländervergleich. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K. Tillmann & M. Weiß (Hrsg.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 211–242). Opladen: Leske und Budrich.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200–230). Münster: Waxmann.
- Stangor, C., Carr, C. & Kiang, L. (1998). Activating Stereotypes Undermines Task Performance Expectations. *Journal of Personality and Social Psychology*, 75, 1191–1197.

- Statistisches Bundesamt (2014a). *Bildung und Kultur: Allgemeinbildende Schulen im Schuljahr 2013/14*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2014b). *Bildungsstand der Bevölkerung 2013*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2014c). *Schulen auf einen Blick*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2014d). *Bildung und Kultur. Schnellmeldungsergebnisse zu Schülerinnen und Schülern der allgemeinbildenden und beruflichen Schulen*. Wiesbaden: Statistisches Bundesamt.
- Stauber, B. (2012). Jugendkulturelle Selbstinszenierungen und (geschlechter-) biographische Relevanzen. In J. Ecarius & M. Eulenbach, *Jugend und Differenz* (S. 51–73). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613–629.
- Steele, C. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology*, 21, 261–302.
- Steele, C. & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811.
- Steele, C., Spencer, S. & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379–440.
- Steele, J. & Ambady, N. (2006). "Math is Hard!" The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology*, 42, 428–436.
- Steffens, M. & Jelenec, P. (2011). Separating Implicit Gender Stereotypes regarding Math and Language: Implicit Ability Stereotypes are Self-serving for Boys and Men, but not for Girls and Women. *Sex Roles*, 64, 324–335.
- Steffens, M., Jelenec, P. & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, 102, 947–963.
- Steinmayr, R. & Spinath, B. (2008). Sex differences in school achievement: what are the roles of personality and achievement motivation? *European Journal of Personality*, 22, 185–209.
- Steins, G. (2005). *Sozialpsychologie des Schulalltags*. Stuttgart: Kohlhammer.
- Stipek, D. & Gralinski, J. (1991). Gender Differences in Children's Achievement-Related Beliefs and Emotional Responses to Success and Failure in Mathematics. *Journal of Educational Psychology*, 83, 361–371.
- Stoet, G. & Geary, D. (2012). Can Stereotype Threat Explain the Gender Gap in Mathematics Performance and Achievement? *Review of General Psychology*, 16, 93–102.
- Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin*, 28, 1667–1678.
- Strack, F. & Deutsch, R. (2002). Urteilsheuristiken. In Frey, D. (Hrsg.), *Theorien der Sozialpsychologie, Band 3: Motivation und Informationsverarbeitung* (S. 352–384). Bern: Huber.
- Strahan, E., Wilson, A., Cressman, K. & Buote, V. (2006). Comparing to perfection: How cultural norms for appearance affect social comparisons and self-image. *Body Image*, 3, 211–227.
- Stricker, L. & Ward, W. (2004). Stereotype Threat, inquiring About Test Takers' Ethnicity and Gender, and Standardized Test Performance. *Journal of Applied Social Psychology*, 34, 665–693.
- Stroessner, S. & Good, C. (2011). *Stereotype Threat: An Overview*. Zugriff am 08.07.2014 von: <http://128.196.134.37/sites/arizona.edu/files/users/user14/Stereotype%20Threat%20Overview.pdf>

- Struch, N. & Schwartz, S. (1989). Intergroup Aggression: Its Predictors and Distinctness. *Journal of Personality and Social Psychology*, 56, 364–373.
- Suls, J., Martin, R. & Wheeler, L. (2002). Social Comparison: Why, With Whom, and With What Effect? *Current Directions in Psychological Science*, 11, 159–163.
- Suls, J. & Wheeler, L. (2000). *Handbook of Social Comparison: Theory and Research*: Springer.
- Suls, J. & Wills, T. (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, 33, 1–39.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Oxford: Academic Press.
- Tajfel, H. (1972). Experiments in a vacuum. In J. Israel & H. Tajfel (Hrsg.), *The context of social psychology: A critical assessment* (S. 69–121). Oxford: Academic Press.
- Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American*, 223, 96–102.
- Tajfel, H., Billig, M., Bundy, R. & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1, 149–178.
- Tajfel, H. & Turner, J. (1986). The Social Identity Theory of Intergroup Behavior. In S. Worchel & W. Austin (Hrsg.), *Psychology of intergroup relations* (S. 7–24). Chicago, IL: Nelson-Hall.
- Tajfel, H. & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel, *The social psychology of intergroup relations* (S. 33–47). Monterey, CA: Brooks & Cole.
- Tarrant, M. (2002). Adolescent Peer Groups and Social Identity. *Social Development*, 11, 110–123.
- Taylor, S., Aspinwall, L., Giuliano, T. & Reardon, K. (1993). Storytelling and Coping with Stressful Events. *Journal of Applied Social Psychology*, 23, 703–733.
- Taylor, S. & Lobel, M. (1989). Social Comparison Activity under Threat: Downward Evaluation and Upward Contacts. *Psychological Review*, 96, 569–575.
- Taylor, S., Wayment, H. & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. Sorrentino & E. Higgins (Hrsg.), *Handbook of motivation and cognition. Foundations of social behavior* (S. 3–27). New York: Guilford Press.
- Tenenbaum, H. & Leaper, C. (2002). Are parents' gender schemas related to their children's gender-related cognitions? A meta-analysis. *Developmental Psychology*, 38, 615–630.
- Terry, D. & Callan, V. (1998). In-Group Bias in Response to an Organizational Merger. *Group Dynamics: Theory, Research, and Practice*, 2, 67–81.
- Terry, D. & O'Brien, A. (2001). Status, legitimacy, and ingroup bias in the context of an organizational merger. *Group Processes & Intergroup Relations*, 4, 271–289.
- Tesser, A., Millar, M. & Moore, J. (1988). Some Affective Consequences of Social Comparison and Processes: the Pain and Pleasure of Being Close. *Journal of Personality and Social Psychology*, 54, 49–61.
- Thoman, D., Smith, J., Brown, E., Chase, J. & Lee, J. (2013). Beyond Performance: A Motivational Experiences Model of Stereotype Threat. *Educational Psychology Review*, 25, 211–243.
- Townsend, S., Major, B., Gangi, C. & Mendes, W. (2011). From "In the Air" to "Under the Skin": Cortisol Responses to Social Identity Threat. *Personality and Social Psychology Bulletin*, 37, 151–164.
- Trautner, H., Ruble, D., Cyphers, L., Kirsten, B., Behrendt, R. & Hartmann, P. (2005). Rigidity and Flexibility of Gender Stereotypes in Childhood: Developmental or Differential? *Infant and Child Development*, 14, 365–381.

- Treumann, K., Sander, U., Meister, D., Strotmann, M., Burkatzki, E., Hagedorn, J. et al. (2007). *Medienhandeln Jugendlicher, Mediennutzung und Medienkompetenz: Bielefelder Medienkompetenzmodell*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Turner, J. (1999). Some Current Issues in Research on Social Identity and Self-categorization Theories. In N. Ellemers, R. Spears & B. Dossje, *Social Identity* (S. 6–34). Oxford: Blackwell.
- Turner, J. & Brown, R. (1978). Social status, cognitive alternatives and intergroup relations. In H. Tajfel (Hrsg.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (S. 201–234). Oxford: Academic Press.
- Turner, J., Hogg, M., Oakes, P., Reicher, S. & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: B. Blackwell.
- Turner, J. & Oakes, P. (1997). The socially structured mind. In C. McGarty & S. Haslam (Hrsg.), *The Message of Social Psychology. Perspectives on Mind in Society* (S. 355–373). Cambridge, MA: Blackwell.
- Turner, J. & Reynolds, K. (2010). The story of social identity. In T. Postmes & N. Branscombe (Hrsg.), *Rediscovering Social Identity: Core Sources* (S. 13–32). Hove: Psychology Press.
- Turner, J. & Reynolds, K. (2001). The Social Identity Perspective in Intergroup Relations: Theories, Themes, and Controversies. In R. Brown & S. Gaertner (Hrsg.), *Blackwell Handbook of Social Psychology: Intergroup processes* (S. 133–152). Malden, MA: Blackwell.
- Turner, R. (1978). Effects of Differential Request Procedures and Self-Consciousness on Trait Attributions. *Journal of Research in Personality*, 12, 431–438.
- Twenge, J. (2001). Changes in women's assertiveness in response to status and roles: A cross-temporal meta-analysis. *Journal of Personality and Social Psychology*, 81, 133–145.
- Twenge, J. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, 36, 305–325.
- Urhahne, D. (2008). Sieben Arten der Lernmotivation: Ein Überblick über zentrale Forschungskonzepte. *Psychologische Rundschau*, 59, 150–166.
- Utman, C. (1997). Performance Effects of Motivational State: A Meta-Analysis. *Personality and Social Psychology Review*, 1, 170–182.
- Van Houtte, M. (2004). Why boys achieve less at school than girls: the difference between boys' and girls' academic culture. *Educational Studies*, 30, 159–173.
- Van Knippenberg, D. (2000). Work Motivation and Performance: A Social Identity Perspective. *Applied Psychology: An International Review*, 49, 357–371.
- Van Knippenberg, A. (1978). Status differences, comparative relevance and intergroup differentiation. In H. Tajfel (Hrsg.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (S. 171–199). London: Academic Press.
- Van Knippenberg, D., Van Dick, R. & Tavares, S. (2007). Social Identity and Social Exchange: Identification, Support, and Withdrawal from the Job. *Journal of Applied Social Psychology*, 37, 457–477.
- Van Knippenberg, A. & van Oers, H. (1984). Social identity and equity concerns in intergroup perceptions. *British Journal of Social Psychology*, 23, 351–361.
- Van Loo, K., Boucher, K., Rydell, R. & Rydell, M. (2013). Competition in stereotyped domains: Competition, social comparison, and stereotype threat. *European Journal of Social Psychology*, 43, 648–660.
- Verkuyten, M. & Nekuee, S. (1999). Ingroup bias: the effect of self-stereotyping, identification and group threat. *European Journal of Social Psychology*, 29, 411–418.
- Vick, S., Seery, M., Blascovich, J. & Weisbuch, M. (2008). The effect of gender stereotype activation on challenge and threat motivational states. *Journal of Experimental Social Psychology*, 44, 624–630.

- Vorauer, J., Hunter, A., Main, K. & Roy, S. (2000). Meta-Stereotype Activation: Evidence from Indirect Measures for Specific Evaluative Concerns Experienced by Members of Dominant Groups in Intergroup Interaction. *Journal of Personality and Social Psychology*, 78, 690–707.
- Vorauer, J. & Kumhyr, S. (2001). Is this about You or Me? Self-Versus Other-Directed Judgments and Feelings in Response to Intergroup Interaction. *Personality and Social Psychology Bulletin*, 27, 706–719.
- Vorauer, J., Main, K. & O'Connell, G. (1998). How Do Individuals Expect to Be Viewed by Members of Lower Status Groups? Content and Implications of Meta-Stereotypes. *Journal of Personality and Social Psychology*, 75, 917–937.
- Voyer, D. & Voyer, S. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204.
- Walton, G. & Cohen, G. (2007). A Question of Belonging: Race, Social Fit, and Achievement. *Journal of Personality and Social Psychology*, 92, 82–96.
- Walton, G. & Cohen, G. (2003). Stereotype Lift. *Journal of Experimental Social Psychology*, 39, 456–467.
- Warrington, M., Younger, M. & Williams, J. (2000). Student Attitudes, Image and the Gender Gap. *British Educational Research Journal*, 26, 393–407.
- Weaver, J., Vandello, J. & Bosson, J. (2013). Intrepid, imprudent, or impetuous? The effects of gender threats on men's financial decisions. *Psychology of Men & Masculinity*, 14, 184–191.
- Weaver-Hightower, M. (2003). The "Boy Turn" in Research on Gender and Education. *Review of Educational Research*, 73, 471–498.
- Weiner, B. (1986). *An Attributional Theory of Motivation and Emotion*. New York, NY: Springer.
- Werner, R. & Collani, G. (2002). *Eine deutsche Skala zu ambivalent-sexistischen Einstellungen (Hostilität und Benevolenz) gegenüber Männern (ASEM)*. Braunschweig: Unveröffentlichtes Manuskript.
- West, C. & Zimmerman, D. (1987). Doing Gender. *Gender & Society*, 1, 125–151.
- Wheeler, L. & Miyake, K. (1992). Social Comparison in Everyday Life. *Journal of Personality and Social Psychology*, 62, 760–773.
- Willingham, W. & Cole, N. (1997). *Gender and fair assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wills, T. (1991). Similarity and self-esteem in downward comparison. In J. Suls & T. Wills (Hrsg.), *Social comparison: Contemporary theory and research* (S. 51–78). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wills, T. (1981). Downward Comparison Principles in Social Psychology. *Psychological Bulletin*, 90, 245–271.
- Wilson, T. & Linville, P. (1985). Improving the Performance of College Freshmen with Attributional Techniques. *Journal of Personality and Social Psychology*, 49, 287–293.
- Winters, M., Haight, R., Swaim, T. & Pickering, K. (2013). The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data. *Economics of Education Review*, 34, 69–75.
- Wissenschaftliches Institut der Ärzte Deutschlands (WIAD) (2000). *Bewegungsstatus von Kindern und Jugendlichen in Deutschland*. Bonn: Forschungsbericht im Auftrag des DSB und der AOK.
- Wood, J., Taylor, S. & Lichtman, R. (1985). Social Comparison in Adjustment to Breast Cancer. *Journal of Personality and Social Psychology*, 49, 1169–1183.
- Wood, W. & Eagly, A. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, 128, 699–727.

- Woolf, K., McManus, C., Gill, D. & Dacre, J. (2009). The effect of a brief social intervention on the examination results of UK medical students: a cluster randomised controlled trial. *BMC Medical Education*, 9, 35–49.
- Wout, D., Danso, H., Jackson, J. & Spencer, S. (2008). The many faces of stereotype threat: Group-and self-threat. *Journal of Experimental Social Psychology*, 44, 792–799.
- Wraga, M., Duncan, L., Jacobs, E., Helt, M. & Church, J. (2006). Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychonomic Bulletin & Review*, 13, 813–819.
- Wraga, M., Helt, M., Jacobs, E. & Sullivan, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social Cognitive and Affective Neuroscience*, 2, 12–19.
- Zagefka, H. & Brown, R. (2005). Comparisons and perceived deprivation in ethnic minority settings. *Personality and Social Psychology Bulletin*, 31, 467–482.
- Zemore, S., Fiske, S. & Kim, H. (2000). Gender Stereotypes and the Dynamics of Social Interaction. In T. Eckes & H. Trautner (Hrsg.), *The Developmental Social Psychology of Gender* (S. 207–241). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zimmer, K., Burba, D. & Rost, J. (2004). Kompetenzen von Jungen und Mädchen. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs* (S. 211–223). Münster: Waxmann.

PUBLIKATIONSVERZEICHNIS

Kessels, U., Heyder, A., Latsch, M. & Hannover, B. (2014). How Gender Differences in Academic Engagement are Related to Students' Gender Identity. *Educational Research, 56*, 220-229.

Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!? How the discourse on "failing boys" impacts performances and motivational goal orientation in German school students. *Social Psychology, 45*, 112-126.

Latsch, M. & Hannover, B. (under review). Ambivalent sexism towards teenage girls. Variations in adolescents' hostile and benevolent sexism depending on threats to or reaffirmations of males' superior social status. Manuscript submitted to *Sex Roles*.

ERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht verwendet. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, Juli 2014

(Martin Latsch)

APPENDIX

Anhang A

Tabelle 1: Übersicht zu bisher veröffentlichten Studien zum Einfluss von *stereotype threat* auf Leistung

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Ambady, N., Paik, S., Steele, J., Owen-Smith, A. & Mitchell, J. (2004)	1 von 2	Frauen in Mathematik	40 Frauen	ja	Gender-primed (weibliche Begriffe), aber individuierte (Fragen zu ihrer Person) Frauen signifikant bessere Leistungen als nicht-individuierte, gender-primed Frauen und ungefähr genauso gut wie unprimed, nicht-individuierte Frauen
Ambady, N., Paik, S., Steele, J., Owen-Smith, A. & Mitchell, J. (2004)	2 von 2	Frauen in Mathematik	39 Frauen	ja	Gender-primed (weibliche Begriffe), aber individuierte (positive und negative Merkmale zur eigenen Person) Frauen signifikant bessere Leistungen als nicht-individuierte, gender-primed Frauen
Aramovich, N. (2014)	1 von 1	Frauen in Mathematik	171 Frauen	ja	Frauen in Gruppen zeigen keine durch stereotype threat beeinflusste Leistung, bei individueller Testung schneiden Frauen unter stereotype threat allerdings schlechter ab, als die Kontrollgruppe.
Armenta, B. (2010)	1 von 1	Ethnische Herkunft in Mathematik	52 Männer & Frauen	ja	Nach Aktivierung des Stereotyps von mathematikbegabten Asiaten verbesserte Leistung der asiatischen und verschlechterte Leistung der hispanischen Stichprobe im Vergleich zur Kontrollgruppe
Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C. & Brown, J. (1999)	2 von 2	Weißer Männer in Mathematik	49 Männer	ja	Männer, die sich stark mit Mathematik identifizieren, unter stereotype threat schlechtere Mathematikleistung als in der KG und als Männer, die sich nicht so sehr mit Mathematik identifizieren.
Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C. & Brown, J. (1999)	1 von 2	Weißer Männer in Mathematik	23 Männer	ja	Weißer Männer unter stereotype threat signifikant ($p < .01$) schlechtere Matheleistung als in der KG
Bailey, A. (2004)	1 von 1	Frauen in Mathematik	143 Männer & Frauen	ja	Frauen unter stereotype threat schlechtere Leistung als Männer und als KG und „female benefit“-Gruppe. Kein signifikanter Unterschied zwischen Männer und Frauen in KG und „female benefit“-Gruppe
Beaton, A., Tougas, F., Rinfret, N. & Huard, N. (2007)	1 & 2 von 2	Frauen und Mathematik	99 Frauen	ja	Frauen nach stereotype threat verminderte Testleistung in Mathematik, insbesondere bei alleiniger Bearbeitung des Tests. In einer Gruppensituation liegen derartige Effekte nur in stark abgeschwächter Form vor.
Bell, A., Spencer, S., Iserman, E. & Logel, C. (2003)	1 von 1	Frauen in Ingenieurwissenschaften	k.A. (nur Abstract verfügbar)	ja	Unter stereotype threat verminderte Testleistung bei Frauen in ingenieurwiss. Test (im Vergleich zu Männern). Abschwächung des Stereotyps verbessert Testleistung auf Niveau der KG.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Brodish, A. & Devine, P. (2009)	1 & 2 von 2	Frauen und Mathematik	je 47 Frauen	ja	Unter stereotype threat verminderte mathematische Testleistung bei Frauen, die sich alle sehr mit Mathematik und dem eigenen Geschlecht identifizieren.
Brown, R. & Day, E. (2006)	1 von 1	Afro-Amerikaner, kognitive Fähigkeiten	136 Männer & Frauen	ja	Afro-Amerikaner unter low stereotype (einige Rätsel) bessere Leistungen als unter standard (Test zum Beobachten und klaren Denken) oder high (IQ-Test) stereotype
Brown, R. & Josephs, R. (1999)	1 von 3	Frauen in Mathematik	126 Männer & Frauen	ja	Frauen, die denken, einen Test durchzuführen, der zeigt, ob sie besonders schwach sind schlechtere Matheleistung als beim Test, der zeigen soll, ob sie besonders stark sind. Umgekehrt bei Männern. In beiden Konditionen schlechtere Leistung der Frauen
Brown, R. & Josephs, R. (1999)	2 von 3	Frauen in Mathematik	74 Männer & Frauen	ja	Frauen mit Handicap (vor dem Test keine Möglichkeit zur Aufwärmung) bei Test, der zeigen soll, ob man besonders schlecht ist, bessere Mathematikleistungen als ohne Handicap. Bei Männern kein Unterschied. Frauen mit Handicap besser Leistung als Männer mit und ohne Handicap
Brown, R. & Pinel, E. (2003)	1 von 1	Frauen in Mathematik	49 Frauen	ja	Frauen unter stereotype threat schlechtere Leistungen als in KG
Cadinu, M., Maas, A., Rosabianca, A., Lombardo, M. & Figerio, S. (2006)	1 & 2 von 2	Frauen mathematische Intelligenz	78 Männer & Frauen	ja	Ergebnisse der Stereotypenbedrohung für beide Geschlechter hochgradig abhängig vom locus of control. Unter internal locus of control schlechteste Leistung beider Geschlechter unter stereotype threat, mit external locus of control keine nennenswerten Unterschiede zwischen den Untersuchungsbedingungen.
Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L. & Latinotti, S. (2003)	1 von 2	Männer und soziale Intelligenz Frauen in Mathematik	95 Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als in KG oder der Gruppe mit positiver Information (Frauen = bessere Leistungen als Männer)
Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L. & Latinotti, S. (2003)	2 von 2	Afro-Amerikaner im Sprachtest	100 Afro-Amerikaner	nein	Mitglieder beider Konditionen, Minderheit (schwarz) und Mehrheit (amerikanisch,) erzielten schlechtere Ergebnisse im Sprachtest
Cadinu, M., Maass, A., Rosabianca, A. & Kiesner, J. (2005)	1 von 1	Frauen in Mathematik	60 Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen und negativere Gedanken über den Mathematiktest
Campbell, S. & Collaer, L. (2009)	1 von 2	Frauen und visuell-räumliches Denken	70 Männer & Frauen	ja	Frauen nach explizitem stereotype threat verminderte Testleistung in visuell-räumlichem Test im Vergleich zu Männern und zu Frauen der Kontrollgruppe, NUR bei schweren Aufgaben

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Campbell, S. & Collaer, L. (2009)	2 von 2	Frauen und visuell-räumliches Denken	70 Männer & Frauen	ja	Frauen nach implizitem stereotype threat verminderte Testleistung in visuell-räumlichem Test im Vergleich zu Männern und zu Frauen der Kontrollgruppe, allerdings NUR bei schweren Aufgaben
Clark, J., Eno, C. & Guadagno, R.	1 von 4	Akad. Leistung von U.S.-Südstaatlern	47 Männer & Frauen	ja	Nach Aktivierung der Herkunft und Bedrohung durch leistungsbezogenes Stereotyp verminderte Leistungsfähigkeit in mathematischem und verbalem Test im Vergleich zur Kontrollgruppe.
Clark, J., Eno, C. & Guadagno, R.	2 von 4	Akad. Leistung von U.S.-Südstaatlern	114 Männer & Frauen	ja	Replikation der Befunde aus Studie 1.
Cohen, G. & Garcia, J. (2005)	2 von 3	Afro-Amerikaner, kognitive Fähigkeiten	57 Männer & Frauen	ja	Schwarze unter stereotype threat weniger Selbstbewusstsein und schlechtere Leistung als in beiden KG's
Crisp, R., Bache, L. & Maitner, A. (2009)	1 & 2 von 2	Frauen und Mathematik	78 Frauen (St. 1) 80 Frauen (St. 2)	nein (Pre-Post-Design)	Psychologiestudentinnen erzielten schlechtere Leistungen in einem Mathematiktest, wenn ihre Gruppenzugehörigkeit aktiviert wurde und der Test als sensibel für Geschlechterunterschiede beschrieben wird. Studentinnen der Ingenieurwissenschaften zeigen hingegen eine verbesserte Testleistung. Studie 2 repliziert diese Befunde.
Croizet, J., Despres, G., Gauzins, M., Huguet, P., Leyens, J. & Méot, A. (2004).	1 von 1	Intelligenz zweier Studierenden-gruppen	109 Männer & Frauen	ja	Durch Labelling eines Raven-Matrizen-tests als diagnostisch für logisches Schlussfolgern und allg. Intelligenz lässt die Testleistung von Studierenden der Psychologie nach, während Studierende der Naturwissenschaft eine verbesserte Testleistung zeigen (jeweils relativ zur Kontrollgruppe)
Davies, P., Spencer, S., Quinn, D. & Gerhardstein, R. (2002)	1 von 3	Frauen in Mathematik	54 Männer & Frauen	ja	Frauen unter stereotype threat (stereotypische TV-Werbungen gesehen) schlechtere Mathematikleistungen als Männer und Frauen in KG
Davies, P., Spencer, S., Quinn, D. & Gerhardstein, R. (2002)	2 von 3	Frauen in Mathematik	67 Männer & Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen, Vermeidung von Mathematikaufgabe und Bevorzugung von sprachlichen Aufgaben als Männer und Frauen in KG
Davies, P., Spencer, S., Quinn, D. & Gerhardstein, R. (2002)	3 von 3	akademische und berufliche Ziele von Frauen	62 Männer & Frauen	ja	Frauen unter stereotype threat weniger Interesse an Berufsfeldern, in denen Mathematik gefordert wird (quantitativer Bereich), und mehr Interesse an Berufsfeldern, in denen sprachliche Kompetenz gefordert wird (sprachlicher Bereich), als Männer und Frauen und KG

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Deemer, E., Thoman, D., Chase, J. & Smith, J. (2014)	1 von 1	Frauen in Naturwissenschaft	493 Frauen	nein	Einfluss von stereotype threat auf Karrierewege in Physik, jedoch nicht in Chemie, vermittelt über naturwissenschaftliches Selbstkonzept.
Dunst, B., Benedek, M., Bergner, S., Athenstaedt, U. & Neubauer, A. (2013)	1 von 1	Frauen und mentale Rotation	64 Jungen & Mädchen	ja	Keine Einflüsse einer Stereotypenbedrohung auf die mentale Rotationsleistung von Mädchen.
Eckert, C. & Imhoff, M. (2013)	1 von 1	Jungen und Leseleistung	222 Jungen	(nein)	Jungen unter stereotype threat erzielen keine geringeren Leseleistungen als Kontrollprobanden.
Elizaga, R. & Markman, K. (2008)	1 von 1	Frauen in Mathematik	158 Männer & Frauen	ja	Frauen mit Partner der in-group (Frau mit schlechten Mathematikleistungen) oder der out-group (Mann mit guten Mathematikleistungen) schlechtere Mathematikleistungen unter diagnostischer Bedingung als in nicht-diagnostischer Bedingung
Eriksson, K. & Lindholm, T. (2007)	1 von 1	Frauen und Mathematik	186 Männer & Frauen	ja	Frauen erzielen nach der experimentellen Erhöhung der Relevanz von Geschlecht schlechtere Testleistungen in einem standardisierten Mathematiktest, allerdings nur wenn diese Geschlecht nicht als zentralen Bestandteil des Selbst ansehen (low identifiers). High identifiers ohne Befund.
Ford, T., Ferguson, M., Brooks, J. & Hagadone, K. (2004)	2 von 2	Frauen in Mathematik	31 Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als in KG und als Frauen, die humorvoll mit Problemen umgehen (coping sense of humour)
Franceschini, G., Galli, S., Chiesi, F. & Primi, C. (2014)	1 von 1	Frauen in Mathematik	78 Frauen	nein	Nur Frauen mit starken, impliziten Stereotypen vulnerable gegenüber Stereotypenbedrohung (massiv geschwächte Leistung), Frauen mit schwachen, impliziten Stereotypen weitestgehend unaffektiert durch stereotype threat
Galdi, S., Cadinu, M. & Tomasetto, C. (2014)	1 von 1	Mädchen in Mathematik	240 Jungen & Mädchen	ja	Nachweis von impliziten, geschlechterbezogenen Mathematik-Stereotypen, welche in einer stereotypenkonsistenten Bedingung schlechtere Mathematikleistung von Schülerinnen verursachen.
Ganley, C., Mingle, L., Ryan, A., Ryan, K., Vasilyeva, M. & Perry, M. (2013)	1 – 3 von 3	Mädchen in Mathematik	931 Jungen & Mädchen	ja	Variation der Aktivierung von stereotype threat (implizit bis explizit), aber kein systematischer Einfluss auf die mathematische Testleistung von Mädchen.
Gibson, C., Losee, J. & Vitiello, C. (2014)	1 von 1	Asiatische Frauen in Mathematik	164 asiatische Frauen	ja	Priming auf asiatische Identität erhöhte im Vergleich zur KG die Mathematikleistung, Priming auf weibliche Identität verminderte die Mathematikleistung bedeutsam.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Gonzales, P., Blanton, H. & Williams, K. (2002)	1 von 1	Frauen und ethnische Herkunft und Mathematik (double minority)	120 Männer & Frauen	(ja)	Darstellung eines mathematischen Tests als diagnostisch für aktuelle Fähigkeiten führt bei Frauen der ethnischen Minderheit (Latinas) zu schlechtester Testleistung, gefolgt von Latinos und weißen Frauen. Ohne dieses Labelling keinerlei Leistungsunterschiede.
Good, C., Aronson, J. & Harder, J. (2008)	1 von 1	Frauen und Mathematik	174 Männer & Frauen (STEM-Studierende)	ja	Stereotype nullification in der Experimentalbedingung konnte die Leistung von Frauen bedeutsam steigern (über das Leistungsniveau der Männer). Ohne weitere Informationen über den Test und Geschlechterunterschiede unterscheiden sich die Leistungen von Frauen und Männern nicht bedeutsam.
Gresky, D., Eyck, L., Lord, C. & McIntyre, R. (2005)	1 von 1	Frauen in Mathematik	94 Frauen & 35 Männer	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als in KG und als Männer
Hartley, B. & Sutton, R. (2013)	2 von 3	Jungen und akad. Leistung (generell)	162 Jungen und Mädchen	ja	Jungen nach Bedrohung als generell akademisch unterlegen zeigen schlechtere Leistung in allen untersuchten Testbereichen (Lesen, Schreiben, Mathematik).
Hirnstain, M., Freund, N. & Hausmann, N. (2012)	1 von 1	Männer und verbale Testleistung	85 Männer & Frauen	ja	Höhere Leistung von Männern nach dem Hinweis auf Geschlechterunterschiede im Wortflüssigkeitstest im Vergleich zu Männern der Kontrollgruppe. Generell höhere Leistung von Frauen, dort aber ebenfalls bessere Testleistung nach Hinweis auf Geschlechterunterschieden zugunsten von Frauen.
Hively, K. & El-Alayli, A. (2014)	1 von 2	Frauen in Sport	61 Männer & Frauen	ja	Athletische Leistung bei Frauen nur in schwieriger Aufgabe vermindert, in leichter Aufgabe kein Effekt
Huguet, P. & Regner, I. (2007)	1 von 2	Mädchen in Mathematik	40 Jungen & Mädchen	ja	Darstellung eines mathematischen Test als „Gedächtnisspiel“ führt zu besserer Testleistung von Mädchen (im Vergleich zu Jungen), wohingegen die Darstellung als geometrische Aufgabe mit verminderter Leistung von Mädchen einhergeht.
Huguet, P. & Regner, I. (2007)	2 von 2	Mädchen in Mathematik	454 Jungen & Mädchen	ja	Replikation der Ergebnisse aus Studie 1. Zusätzlich konnte gezeigt werden, dass Mädchen bei der Bearbeitung der Aufgaben (auch und insbesondere in der Bedingung, in der die Aufgabe als geometrisch/mathematisch dargestellt wird) in gleichgeschlechtlichen Gruppen enorm profitieren (Anhebung des generellen Leistungsniveaus, Minimierung von Geschlechtsunterschieden).
Joanisse, M., Gagnon, S. & Voloaca, M. (2013)	1 von 1	Alter und Fahrleistung	61 Männer & Frauen	ja	Stereotype threat beeinträchtigt die Fahrleistung von älteren Probanden im Vergleich zur Kontrollgruppe.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
John-Henderson, N., Rheinschmidt, M., Mendoza-Denton, R. & Francis, D. (2014)	1 von 2	Probanden mit geringem SES	90 Männer & Frauen	ja	Probanden mit geringem SES schnitten in einem Test zu verbalen Fähigkeiten in der threat-Bedingung schlechter ab, als die Kontrollgruppe
John-Henderson, N., Rheinschmidt, M., Mendoza-Denton, R. & Francis, D. (2014)	2 von 2	Probanden mit geringem SES	98 Männer & Frauen	ja	Einfluss von sozialen Vergleichsprozessen: Probanden mit geringem SES antizipieren im aufwärtsgerichteten Vergleich schlechtere Testleistung
Johns, M., Schmader, T. & Martens, A. (2005)	1 von 1	Frauen in Mathematik	75 Frauen & 42 Männer	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als in KG und als Männer
Josephs, R., Newman, M., Brown, R. & Beer, J. (2003)	1 von 2	Frauen in Mathematik	151 Männer & Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als in KG und als Männer
Josephs, R., Newman, M., Brown, R. & Beer, J. (2003)	2 von 2	Männer in Mathematik	51 Männer	ja	Männer unter stereotype threat bessere Mathematikleistungen als in KG
Keller, J. & Bless, H. (2008)	2 von 3	Frauen im räumlichen Denkvermögen	178 Frauen	ja	Prevention focus aktiviert: negative Erwartungen führten zu schlechteren Ergebnissen als positive Promotion focus aktiviert: Negative Erwartungen führten zu besseren Ergebnissen als positive KG: keine reliablen Unterschiede zwischen positiven und negativen Erwartungen
Keller, J. & Dauenheimer, D. (2003)	1 von 1	Frauen in Mathematik	74 Mädchen & Jungen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen (versuchten auch weniger Aufgaben zu lösen) als Frauen in KG und Männer in EG
Keller, J. (2002)	1 von 1	Mädchen in Mathematik	75 Mädchen & Jungen	ja	Mädchen unter stereotype threat schlechtere Mathematikleistungen als in KG und als Jungen.
Keller, J. (2007)	2 von 2	Männer und verbale Testleistung	59 Männer	(nein)	Abhängigkeit der Ergebnisse von stereotype threat von regulatorischen Foki: nach negativem threat unter einem prevention focus verringerte verbale Testleistung im Vergleich zu Männern mit aktiviertem promotion focus.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Koch, S., Konigorski, S. & Sieverding, M. (2014)	1 & 2 von 2	Frauen in Deutsch und Mathematik	92 Männer & Frauen	ja	Einfluss von Sexismus des Versuchsleiters als stereotype threat auf die Mathematikleistung von Frauen, jedoch nicht auf die Deutschleistung (im Vergleich zur KG).
Latsch, M. & Hannover, B. (2014)	2 von 4	Jungen in Deutsch und Mathematik	124 Jungen & Mädchen	ja	Einfluss einer generellen Stereotypenbedrohung bei Jungen nur in Deutsch, keine Effekte in Mathematik → Verstärkung Geschlechterstereotype
Laurin, R. (2013)	1 von 1	Frauen und Mathematik	161 Männer & Frauen	ja	Je nach Anlage der Bedrohung (Nachteile für Frauen, Nachteile für Männer, keine Geschlechterunterschiede) stereotype lift für Frauen oder für Männer (wenn entsprechende Bedingung), stereotype threat allerdings nur bei Frauen.
Lesko, A. & Corpus, J. (2006)	1 von 1	Frauen und Mathematik	128 Männer & Frauen	(ja)	Frauen unter stereotype threat schlechtere Testleistung in Mathematik als Männer oder Frauen in einer stereotype nullification Bedingung.
Martens, A., Johns, M., Greenberg, J. & Schimel, J. (2006)	1 von 2	Frauen in Mathematik	59 Männer & Frauen	ja	Frauen unter stereotype threat schlechtere Mathematikleistungen als Frauen in KG und als Männer Frauen unter stereotype threat und vorheriger self-affirmation bessere Leistungen als Frauen nur unter stereotype threat
Martens, A., Johns, M., Greenberg, J. & Schimel, J. (2006)	2 von 2	Frauen in räumlichem Denken	105 Männer & Frauen	nein	Frauen unter stereotype threat <u>ohne</u> vorherige self-affirmation erzielten schlechtere Ergebnisse als Frauen unter stereotype threat <u>mit</u> self-affirmation Bei Männern (die ja nicht unter stereotype threat leiden) kein signifikanter Unterschied Unterschied zwischen Frauen und Männern <u>mit</u> self-affirmation nicht signifikant
Marx, D., Monroe, A., Cole, C. & Gilbert, P. (2013)	1 von 1	Frauen und Mathematik	122 Männer & Frauen	nein	Einfluss von Rollenmodellen (Ingroup vs. Outgroup) und deren Zweifeln an der eigenen mathematischen Fähigkeit. Weibliche Rollenmodelle, die Zweifel an eigener Mathematik-Fähigkeit äußern führen zu Minderleistung von Frauen und höherer Leistung von Männern. Bei Outgroup-Rollenmodellen (Frau als Modell bei Männern und vv.) keinerlei systematischer Einfluss.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
McFarland, L., Lev-Arey, D. & Ziegert, J. (2003)	1 von 1	Afro-Amerikaner, kognitive Fähigkeiten	241 Männer & Frauen	ja	Zusammenhang zw. threat und Leistung hängt sowohl von der ethnischen Identität <u>als auch</u> von der domain identification ab Es zeigte sich ein signifikant positiver Zusammenhang zw. Veränderungen in der ethnischen Identität & der Testleistung im kogn. Fähigkeitstest von Schwarzen, nicht aber von Weißen (Schwarze, deren ethnische Identität abnahm, zeigten bessere Leistungen). In der Threat-Bedingung zeigten signifikant bessere Leistungen im Persönlichkeitstest als in der Kontrollbedingung (→ Threat generalisiert sich & beeinflusst Leistung auch in Tests, für die kein Stereotyp existiert).
McIntyre, R., Lord, C., Gresky, D., Eyck, L., Frye, G. & Bond, C. (2005)	1 von 1	Frauen in Mathematik	295 Männer & Frauen	ja	Stereotype wurden bei allen VPs erwähnt, aber die Anzahl vorgegebener, kurzer Biografien über erfolgreiche Frauen (0-4) variierte. Frauen, die keine Biografien lasen, zeigten schlechtere Leistungen als Männer; Frauen die vier lasen waren gleich gut. Der Leistungsanstieg (der Frauen) über die Anzahl der Rollenmodelle war konsistent mit der von der Social Impact Theory (Latané, 1981) vorhergesagten potenziellen Funktion.
McIntyre, R., Paulson, R. & Lord, C. (2003)	1 von 2	Frauen in Mathematik	162 Männer & Frauen	ja	Frauen zeigten signifikant bessere Leistung in schwierigen Math-Tests, wenn ihnen vorher gesagt wurde, dass Frauen generell bessere Teilnehmer in psychologischen Experimenten seien als Männer.
McIntyre, R., Paulson, R. & Lord, C. (2003)	2 von 2	Frauen in Mathematik	106 Männer & Frauen	ja	Frauen zeigten signifikant bessere Leistung in schwierigen Math-Tests, wenn sie zuvor etwas über vier erfolgreiche Frauen lasen.
Nguyen, H., O'Neal, A. & Ryan, A. (2003)	1 von 1	Afro-Amerikaner, kognitive Fähigkeiten	172 Männer & Frauen	ja	Es zeigten sich signifikante ethnische Unterschiede zugunsten von Weißen in der mathematischen & „logischen“ Leistung, auch wenn die Fähigkeit konstant gehalten wurde. Die Metakognitive Strategie „Kognitionsregulation“ (regulation of cognition; z.B. Planung, wie man den Test effektiv durchführt) medierte den Effekt der Ethnie auf die mathematische & verbale Leistung teilweise. Es wurden jedoch keine signifikanten Stereotype-threat-Effekte gefunden. Es ergab sich aber eine Dreifachinteraktion die zeigte, dass in der Stereotype-threat-Bedingung das Beherrschen von Test-taking-Strategien die Kluft zw. den mathemat. Leistungen der Weißen & Afroamerikaner abmilderte.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
O'Brien, L. & Crandall, C. (2003)	1 von 1	Frauen in Mathematik	164 Männer & Frauen	ja	Frauen in der Gender-differences-Gruppe (es wurde gesagt, dass es in diesem Mathe-Test Geschlechterunterschiede gäbe) zeigten, im Vergleich zur No-differences-Bedingung, bessere Leistungen im einfachen Test und schlechtere Leistungen im schweren Test. Die Leistung der Männer war unabhängig von der Manipulation.
Oswald, D. & Harvey, R. (2000–2001)	1 von 1	Frauen in Mathematik	72 Frauen	nein	Signifikante Interaktion einer frauenfeindlichen Umwelt (hier: abwertender Cartoon über mathemat. Fähigkeiten von Frauen) und stereotype threat. In einer nicht feindlichen Umwelt (kein Cartoon) zeigten Personen der stereotype-threat-remove-Bedingung (Ansage, dass es keinen Unterschied zw. Männern & Frauen gäbe) bessere Leistungen als Personen der stereotype-threat-Bedingung, in der feindlichen Umwelt zeigte sich ein gegenteiliges Muster. Darüber hinaus erlebten Frauen in feindlichen Umwelten und unter stereotype threat verschiedene emotionale Reaktionen (u. A. Frustration, weniger Ruhe, Unsicherheit) während des Tests.
Picho, K. & Stephens, J. (2012)	1 von 1	Frauen und Mathematik	89 Jungen & Mädchen	ja	Mädchen in koedukativen Schulen erzielen nach stereotype threat bedeutsam verringerte Leistungen in einem Mathematiktest im Vergleich zur Kontrollgruppe. In monoedukativen Schulen (nur Mädchen) keinerlei Effekte.
Ployhart, R., Ziegert, J. & McFarland, L. (2003)	1 von 1	Afro-Amerikaner, kognitive Fähigkeiten	394 Männer & Frauen	ja	Untersuchung zum Einfluss von stereotype threat & Wahrnehmung von Bewerbern auf ethn. Subgruppenunterschiede in kogn. Fähigkeitstests in Auswahlkontexten. Über alle stereotype threat und face validity Bedingungen zeigten Schwarze die schlechtesten Leistungen in der diagnostischen Bedingung (Test als Maß für ihre Intelligenz). Es zeigten sich außerdem individuelle Unterschiede in der Wahrnehmung des stereotype threat, welche zu geringerer Augenscheinvalidität, geringerer Testmotivation und höherer Ängstlichkeit beitrugen. Darüber hinaus beeinflusste die Motivation die Testleistung positiv, die Ängstlichkeit beeinflusste diese negativ.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Quinn, D. & Spencer, S. (2001)	2 von 2	Frauen und Mathematik	108 Männer & Frauen	(ja)	Frauen zeigen nach stereotype threat eine geringere Leistung in einem standardisierten Mathematiktest nach expliziter Stereotypenaktivierung. Nach Abschwächung des Stereotyps zeigen Frauen im Vergleich zu Männern stärkere Testleistungen.
Rivardo, M., Rhodes, M., Camaione, T. & Legg, J. (2011)	1 von 1	Frauen und Mathematik	148 Männer & Frauen	ja	Frauen unter stereotype threat zeigen eine verminderte mathematische Testleistung, auch dann wenn alternative Attributionsstrategien für potentiellen Misserfolg vorgegeben wurden.
Rosenthal, H. & Crisp, R. (2006)	2 von 3	Frauen in Mathematik	36 Frauen	nein	Frauen, die über Gemeinsamkeiten von Männern und Frauen nachdenken sollten, zeigten bessere Leistungen verglichen mit der Baseline-Gruppe und verglichen mit Frauen, die über Unterschiede zw. Männern und Frauen nachdenken sollte.
Rosenthal, H. & Crisp, R. (2006)	3 von 3	Frauen in Mathematik	62 Frauen	nein	(Hier wurde stereotype threat <u>explizit</u> induziert.) Frauen, die über Gemeinsamkeiten von Männern und Frauen nachdenken sollten, bevor stereotype threat induziert wurde, zeigten bessere Leistungen verglichen mit Frauen, die danach über die Gemeinsamkeiten nachdenken musste, mit der Baseline-Gruppe und verglichen mit Frauen, bei denen „nur“ stereotype threat induziert wurde (ohne die Gemeinsamkeiten-Aufgabe).
Rydell, R., McConnell, A. & Beilock, S. (2009)	1 von 4	Frauen und Mathematik	112 Frauen	ja	Frauen unter stereotype threat („Frauen sind schlecht in Mathe“) zeigen verminderte Testleistung im Vergleich zur Kontrollgruppe. Nach zusätzlicher Aktivierung einer sozialen Gruppenidentität, über welche positive Leistungsstereotypen vorgehalten werden (Studierende) verschwindet der durch stereotype threat induzierte Effekt vollständig.
Rydell, R., McConnell, A. & Beilock, S. (2009)	2 von 4	Frauen und Mathematik	98 Frauen	ja	Replikation der Ergebnisse aus Studie 1: Aktivierung des Geschlechterstereotyps führt zu verminderter Testleistung in Mathematik, zusätzliche Aktivierung der sozialen Identität als Studierende führt zur Eliminierung negativer threat-Effekte.
Rydell, R., McConnell, A. & Beilock, S. (2009)	3 von 4	Frauen und Mathematik	57 Frauen	ja	Einfluss multipler sozialer Identitäten auf die Arbeitsgedächtniskapazität, welche wiederum stereotype threat-Effekte moderiert.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Rydell, R., McConnell, A. & Beilock, S. (2009)	4 von 4	Frauen und Mathematik	80 Frauen	ja	Nach Angabe des Geschlechts zeigten Frauen eine verminderte Testleistung in Mathematik. Wurden zusätzlich Variablen mit abgefragt, welche eine Mitgliedschaft in einer positiv stereotypisierten sozialen Gruppe beinhalten verschwindet dieser Effekt.
Rydell, R., Van Loo, K. & Boucher, K. (2014)	1 von 3	Frauen in Mathematik	168 Männer & Frauen	ja	Reduzierte Arbeitsgedächtnisleistung unter threat, schlechtere Mathematikleistung als KG oder Männer
Rydell, R., Van Loo, K. & Boucher, K. (2014)	2 von 3	Frauen in Mathematik	90 Frauen	ja	Reduzierte Arbeitsgedächtnisleistung unter threat, schlechtere Mathematikleistung als KG
Rydell, R., Van Loo, K. & Boucher, K. (2014)	3 von 3	Frauen in Mathematik	82 Frauen	ja	Differentielle Einflüsse von stereotype threat auf die Arbeitsgedächtnisfunktion (Verminderung von Updating-Prozessen, Inhibition für Risikowahlverhalten).
Sawyer, T. & Hollis-Sawyer, L. (2005)	1 von 1	Ethnische Herkunft und kognitive Leistungsfähigkeit	189 Männer & Frauen	ja	Interaktion von Persönlichkeitseigenschaften und Testbedingungen (transaktionales Modell) sagt Varianz der Testleistung, Unterschiede in der Prüfungsängstlichkeit und im stereotype threat voraus. Es zeigten sich u.a. signifikante Zusammenhänge zw. Extraversion, Offenheit für Erfahrungen, Verträglichkeit sowie dem Glauben an Intelligenz (fix vs. veränderbar) und stereotype threat. Höher wahrgenommene Level von stereotype threat waren u.a. mit erhöhten <i>stereotype threat belief perceptions</i> , niedrigeren Werten im Kognitiven Fähigkeitstest verbunden.
Schimmel, J., Arndt, J., Banko, K. & Cook, A. (2004)	2 von 3	Frauen in Mathematik	46 Frauen	nein	In der Stereotype-threat-Bedingung erhöhte intrinsische Selbstbestätigung (im Gegensatz zu extrinsischer) die Testleistung.
Schmader, T. (2002)	1 von 1	Frauen in Mathematik	65 Männer & Frauen		Wurde die Geschlechtsidentität zur Leistung im Mathematiktest in Verbindung gebracht, dann zeigten Frauen mit einem höheren Maß an Geschlechtsidentifikation schlechtere Leistungen als Männer (Frauen mit einem niedrigeren Maß an GI zeigten die gleichen Leistungen wie Männer). Wurde die Geschlechtsidentität nicht zur Leistung im Mathematiktest in Verbindung gebracht, dann zeigten Frauen die gleiche Leistung wie Männer unabhängig davon, wie viel Bedeutung sie der Geschlechtsidentität beimaßen.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Schmader, T. & Johns, M. (2003)	1 von 3	Frauen in Mathematik	75 Männer & Frauen	ja	VPn bearbeiteten eine Arbeitsgedächtnis-Aufgabe. In der Stereotype-Threat-Bedingung Labelling als Messung der „quantitative Kapazität“ und dass Geschlechterunterschiede in den mathematischen Fähigkeiten evtl. auf Geschlechterunterschieden in der quantitativen Kapazität basieren. Frauen unter stereotype threat zeigten reduzierte kognitive Kapazität (gemessen anhand Anzahl erinnertes Wörter) im Gegensatz zu Männern und Frauen in Kontroll-Bedingung.
Schmader, T. & Johns, M. (2003)	2 von 3	Ethnische Herkunft und kognitive Leistungsfähigkeit	73 Männer & Frauen	ja	VPn bearbeiteten einen Arbeitsgedächtnis-Test. In der Stereotype-Threat-Bedingung wurde zudem gesagt, dass dieser Test hoch prädiktiv für die Leistung in Intelligenztests sei und mit ihren Werten Normen für verschiedene Gruppen erstellt werden würden (+ sie sollten ihre Ethnie angeben). Latinos unter stereotype threat zeigten reduzierte kognitive Kapazität (gemessen anhand Anzahl erinnertes Wörter) im Gegensatz zu Weißen und Latinos in Kontroll-Bedingung.
Schmader, T. & Johns, M. (2003)	3 von 3	Frauen in Mathematik	31 Frauen	nein	Reduktion der Kapazität des Arbeitsgedächtnisses vermittelt den Effekt des stereotype threat auf die Mathematikleistung von Frauen.
Schmader, T., Johns, M. & Barquissau, M. (2004)	2 von 2	Frauen in Mathematik	68 Frauen	nein	Frauen, die dazu tendierten Geschlechterstereotypen über die mathematischen Fähigkeiten von Frauen zuzustimmen, waren anfälliger für die negativen Effekte von stereotype threat bezüglich ihrer mathematischen Leistungen.
Seibt, B. & Förster, J. (2004)	2 & 3 von 5	Männer und verbale Testleistung	60 Männer & Frauen (St. 2) 56 Männer & Frauen (St. 3)	ja	Männer unter negativem stereotype threat bearbeiten verbale Testaufgaben langsamer, aber genauer – wohingegen bei Vorgabe eines positiven Leistungsstereotyps schneller, aber weniger akkurat (v.a. viele false alarms) reagiert wird.
Sekaquaptewa, D. & Thompson, M. (2002)	1 von 1	Frauen in Mathematik	157 Männer & Frauen	ja	Personen lernten arithmetisches Material und wurden dann in Gruppen des gleichen Geschlechts vs. Gruppen des anderen Geschlechts getestet. Frauen zeigten schlechtere Leistungen bei Testung in Gruppen des anderen Geschlechts und unter stereotype threat (beide Faktoren zusammen war nachträglicher als nur einer). Die Leistung der Männer war über alle Bedingungen hinweg gleich.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Shapiro, J., Williams, A. & Hambarchyan, M. (2013)	2 & 4 von 4	Frauen und Mathematik	68 & 75 Frauen	ja	Frauen weniger Interesse und schlechtere Leistung unter stereotype threat im Vergleich zur KG
Shapiro, J., Williams, A. & Hambarchyan, M. (2013)	1 & 3 von 4	Ethnizität und Intelligenz	75 & 77 Afroamerikaner/ -innen	ja	Afroamerikaner schlechtere Testleistung nach stereotype threat
Smeding, A., Dumas, F., Loose, F. & Régner, I. (2013)	1 von 2	Mädchen in Mathematik	1.127 Jungen & Mädchen	(ja)	Einfluss der Testanordnung auf die mathematische Leistung von Mädchen: wurde zuerst ein verbaler Test durchgeführt ergeben sich im Vergleich zu Jungen keine geringeren, mathematischen Leistungen. Jedoch deutlicher stereotype threat-Effekt, wenn zuerst der mathematische Test vorgelegt wird. Kein stereotype threat-spillover.
Smith, C. & Hopkins, R. (2004)	1 von 1	Afro-Amerikaner, kognitive Fähigkeiten	160 Männer & Frauen	nein	Stereotype threat beeinflusste die Testleistung nicht, aber Schwarze mit hoher kultureller Identität und einem hohen internen Locus of control zeigten signifikant bessere Leistungen.
Smith, J. & White, P. (2002)	1 von 2	Frauen in Mathematik	70 Frauen	nein	Sowohl bei explizit (es wurde gesagt, dass Männer in Mathematik besser seien) als auch bei implizit (keine Anmerkung zu Geschlechterunterschieden) aktivierten stereotype threat zeigten Frauen schlechtere Leistungen als in der Nullified-Bedingung (es wurde gesagt, dass Männer und Frauen gleich gut in Mathematik seien).
Smith, J. & White, P. (2002)	2 von 2	Weißer Männer in Mathematik	73 Männer	nein	Replikation von Experiment 1 mit weißen Männern und dem Stereotyp, das Asiaten besser in Mathematik seien als Weiße. In einer kleinen Feldstudie wurde mittels Wortfragmentaufgaben mit Weißen in Mathematikkursen darüber hinaus gezeigt, dass dieses Stereotyp tatsächlich spontan auftrat.
Spencer, S., Steele, C. & Quinn, D. (1999)	2 von 3	Frauen in Mathematik	54 Männer & Frauen	ja	Frauen unter (hohem) stereotype threat zeigten deutlich schlechtere Leistungen als gleichqualifizierte Männer. Dieser Leistungsunterschied konnte eliminiert werden, wenn gesagt wurde, dass es bei diesem Test keine Geschlechterunterschiede gäbe.
Steele, C. & Aronson, J. (1995)	1 von 4	Afro-Amerikaner, verbaler Fähigkeitstest	117 Männer & Frauen	ja	Schwarze waren in der Diagnostik-Bedingung (es wurde gesagt, dass der Test Maß ihrer intellektuellen Fähigkeit sei) schlechter als Weiße. Dieser Effekt war aber nur sehr schwach. In der Nicht-Diagnostik-Bedingung zeigte sich kein Stereotype threat.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Steele, C. & Aronson, J. (1995)	2 von 4	Afro-Amerikaner, verbaler Test	40 Männer & Frauen	ja	Schwarze waren in der Diagnostik-Bedingung schlechter als Weiße und bearbeiteten auch weniger Items.
Steele, C. & Aronson, J. (1995)	4 von 4	Afro-Amerikaner, verbaler Fähigkeitstest	47 Männer & Frauen	ja	Durch Priming wurde das Stereotyp <u>direkt</u> (im Gegensatz zu Studie 1 und 2, wo es indirekt evoziert wurde) evoziert. Hier war die Testleistung der Schwarzen auch in der Nicht-Diagnostik-Bedingung schlechter.
Stricker, L. & Ward, W. (2004)	1 von 2	Verschiedene Ethnien in Mathematik	1652 Männer & Frauen	ja	Befragung nach Geschlecht und Ethnie hatte keinen signifikanten Effekt auf die Testleistung von Schwarzen, Frauen oder anderen Subgruppen.
Stricker, L. & Ward, W. (2004)	2 von 2	Verschiedene Ethnien in Leistungstest	1341 Männer & Frauen	ja	Befragung nach Geschlecht und Ethnie hatte keinen signifikanten Effekt auf die Testleistung von Schwarzen, Frauen oder anderen Subgruppen.
Taylor, C., Lord, C., McIntyre, R. & Paulson, R. (2011)	1 von 1	Frauen und Mathematik	113 Frauen	ja	Einflüsse von Rollenmodellen auf stereotype threat: Ein nicht-fachbezogenes Rollenmodell verbessert die mathematische Testleistung nach stereotype threat bei Probandinnen (die der Ansicht sind das Rollenmodell verdient seinen Erfolg) auf das Niveau einer Kontrollgruppe ohne threat. Probandinnen, die der Ansicht sind, dass Rollenmodell verdient seinen Erfolg <i>nicht</i> , schneiden vergleichbar schlecht ab wie eine Kontrollgruppe nur unter stereotype threat (ohne Rollenmodell). Positive Rollenmodelle eliminieren somit stereotype threat-Effekte nur, wenn dem Rollenmodell der Erfolg selbst zugeschrieben wird.
Thames, A., Hinkin, C., Byrd, D., Bilder, R., Duff, K., Mindt, M., Arentoft, A. & Streiff, V. (2013)	1 von 1	Frauen und Mathematik	122 Männer & Frauen	ja	Frauen ohne weitere Informationen schneiden in einem Leistungstest unter stereotype threat schlechter ab. Nach dem Hinweis, dass Frauen zwar nicht gleichermaßen wie Männer in STEM-Berufen repräsentiert, aber dennoch in diesem Feld erfolgreich tätig sind Elimination der threat-Effekte.
Thoman, D., White, P., Yamawaki, N. & Koishi, H. (2008)	1 von 1	Frauen und Mathematik	66 Frauen	ja	Unter klassischen stereotype threat Bedingungen (Frauen sind aufgrund von natürlichen Kompetenzunterschieden mathematisch unterlegen) zeigen Frauen eine schlechtere Testleistung als die Kontrollgruppe. Wird jedoch dargestellt, dass Geschlechterunterschiede aufgrund von beeinflussbaren Größen wie Anstrengung und Ausdauer zustande kommen zeigen sich vermehrte Anstrengung und bessere Testleistungen als die Kontrollgruppe.

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Van Loo, K., Boucher, K., Rydell, R. & Rydell, M. (2013)	1 von 4	Frauen und Mathematik	232 Männer & Frauen	ja	Unter Priming auf „Wettbewerb“ zwischen den Geschlechtern schlechtere Testleistung in einem mathematischen Test bei Frauen nach stereotype threat. Nach explizitem Hinweis, dass der Mathematiktest keine Geschlechterunterschiede produziert Eliminierung der Effekte.
Van Loo, K., Boucher, K., Rydell, R. & Rydell, M. (2013)	2 von 4	Frauen und Mathematik	103 Männer & Frauen	ja	Replikation der Ergebnisse auf Studie 1: kompetitiver Prime sorgt für eine verstärkt nachlassende Mathematikleistung bei Frauen (jedoch nicht bei Männern) im Vergleich zur Kontrollgruppe ohne kompetitives Priming.
Van Loo, K., Boucher, K., Rydell, R. & Rydell, M. (2013)	4 von 4	Frauen und Mathematik	65 Männer & Frauen	ja	Männer zeigen nach kompetitivem Prime unter stereotype threat deutlich geringere verbale Testleistungen als ohne kompetitives Priming (Übertragungen der Ergebnisse aus Studien 1 und 2 auf stereotype threat bei Männern).
von Hippel, W., von Hippel, C., Conway, L., Preacher, K., Schooler, J. & Radvansky, G. A. (2005)	4 von 4	Ethnische Herkunft, Intelligenz	56 Männer & Frauen	nein	Weißer, die mit Vorurteil konfrontiert wurden, dass Asiaten besser in diesem Intelligenztest seien, neigten eher dazu zu bestreiten, dass Intelligenz wichtig sei, wenn sie hohes impression management hatten (verwendeten also Leugnung als Impression-Management-Strategie).
Walsh, M., Hickey, C. & Duffy, J. (1999)	2 von 2	Frauen in Mathematik	174 Männer & Frauen	ja	Manipulation von gender labelling (weibl., männl. oder neutraler Charakter) konnte Geschlechterunterschiede nicht erklären, stereotype threat schon.
Weger, U., Hooper, N., Meir, B. & Hopthrow, T. (2012)	1 von 1	Frauen und Mathematik	71 Frauen	ja	Negative Effekte durch stereotype threat bei Frauen in Mathematik (Kontrollgruppe) werden durch ein kurzes Achtsamkeitstraining (Experimentalgruppe) eliminiert.
Wicherts, J., Dolan, C. & Hessen, D. (2005)	1 von 3	Ethnische Minderheiten im Intelligenztest	295 Männer & Frauen	ja	MANOVA zeigte keinen Haupteffekt von stereotype threat (ST) auf die Testleistung der Minderheitengruppe, aber die ST-Manipulation führte zu Messbias bzgl. der Minderheitengruppe.
Wicherts, J., Dolan, C. & Hessen, D. (2005)	3 von 3	Frauen in Mathematik	283 Männer & Frauen	ja	MANOVA: Frauen unter stereotype threat schlechtere Leistungen (in einem Subtest)

Autoren	Studie	stereotyp. Gruppe	N	KG?	Haupt-Ergebnisse
Wout, D., Danso, H., Jackson, J. & Spencer, S. (2008)	1 von 2	Frauen und Mathematik	62 Frauen	ja	Frauen weisen eine geringere mathematische Testleistung in der Bedingung auf, in welcher der Mathematiktest als diagnostisch dargestellt wird (im Vergleich zu einer Kontrollgruppe). Zusätzlich leisten Frauen weniger, wenn sie über ihre eigenen Mathematik-Fähigkeiten nachdenken sollten (self-threat), aber nicht, wenn eine die gesamte soziale Gruppe (group-threat) in ihrer Leistungsfähigkeit bedroht wurde.
Wout, D., Danso, H., Jackson, J. & Spencer, S. (2008)	2 von 2	Frauen und Mathematik	59 Frauen	ja	Einfluss der Identifikation mit dem eigenen Geschlecht: Frauen, die sich hoch mit dem eigenen Geschlecht identifizieren leisten unter einem group-threat weniger, als Frauen mit geringer gender identification. Kein Einfluss der gender identification in der self-threat-Bedingung.

Anhang B: Untersuchungsmaterial

Das Material für die durchgeführten Befragungen einschließlich der Genehmigung durch die Senatsverwaltung für Bildung, Wissenschaft und Forschung kann auf Rückfrage beim Autor angefordert werden: martin.latsch@fu-berlin.de.