

Mehr als Erwartungen und Werte:  
Zur Rolle von Stereotypen für den differenziellen  
Schulerfolg von Mädchen und Jungen

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

am Fachbereich Erziehungswissenschaft und Psychologie  
der Freien Universität Berlin

vorgelegt von

Dipl.-Psych. Anke Heyder

Berlin, 2015



Erstgutachterin: Prof. Dr. Ursula Kessels  
Zweitgutachter: Prof. Dr. Martin Brunner  
Disputation: 13. Mai 2015



Diese Arbeit entstand im Rahmen des von der Deutschen Forschungsgemeinschaft geförderten Projekts „Sind Jungen die neuen Bildungsverlierer?“ unter der Leitung von Prof. Dr. Ursula Kessels (KE 1412/2-1 und KE 1412/2-2). Es handelt sich dabei um ein Gemeinschaftsprojekt in Kooperation mit Prof. Dr. Bettina Hannover.



## **Danksagung**

Diese Arbeit entstand im Rahmen des DFG-Projekts „Sind Jungen die neuen Bildungsverlierer?“. Für die Möglichkeit, in diesem Projekt als Wissenschaftliche Mitarbeiterin arbeiten und meine Dissertation schreiben zu können, bin ich sehr dankbar.

Mein größter Dank gilt hier Ursula Kessels für ihre hervorragende fachliche Unterstützung, ihre Begeisterung, ihren Humor und ihre Hartnäckigkeit, durch die ich in den vergangenen Jahren wissenschaftlich und persönlich viel gelernt habe. Die Zusammenarbeit hat mir viel Spaß gemacht.

Den weiteren Mitgliedern meiner Promotionskommission, Martin Brunner, Bettina Hannover, Rainer Watermann und Lysann Zander, danke ich ganz herzlich für ihr Interesse an meiner Arbeit und ihre Unterstützung bei der Terminfindung. Mein besonderer Dank gilt in diesem Zusammenhang Martin Brunner für die Bereitschaft, das Zweitgutachten zu übernehmen, und Bettina Hannover dafür, dass sie immer für mich ansprechbar war und sich in verschiedensten Angelegenheiten für mich eingesetzt hat.

Außerdem möchte ich mich bei Ricarda Steinmayr bedanken, die vor einigen Jahren die Erste war, die mich ermutigt hat, zu promovieren, und mich seitdem in vielen herausfordernden Situationen fachlich und persönlich unterstützt hat.

Richard Meyer-Eppler, Katarzyna Gogol, Fenne große Deters, Katharina Thoren, Katharina Holder und meiner Familie danke ich für die hilfreichen Rückmeldungen und die Gelassenheit und Zuversicht, mit der sie mir in den letzten Monaten begegnet sind. Für die Hilfe beim Zurechtfinden und Einleben an der FU Berlin bin ich außerdem Martin Latsch sehr dankbar.

Zu guter Letzt möchte ich mich bei allen Kolleginnen und Kollegen und studentischen Hilfskräften in Köln und Berlin bedanken, die mich bei den Datenerhebungen unterstützt haben, sowie bei allen Schülerinnen, Schülern und Lehrkräften, die durch ihre Studienteilnahme diese Arbeit möglich gemacht haben.





## Inhaltsverzeichnis

<b>Zusammenfassung.....</b>	<b>11</b>
<b>Summary.....</b>	<b>15</b>
<b>Einleitung.....</b>	<b>17</b>
<b>1 Geschlechtsunterschiede im Schulerfolg.....</b>	<b>21</b>
1.1 Bildungsbeteiligung und Schulabschlüsse.....	21
1.2 Schulnoten.....	22
1.3 Leistungs- und Kompetenztests.....	26
1.4 Fazit.....	33
<b>2 Geschlechtsunterschiede in individuellen schulerfolgsrelevanten Merkmalen.....</b>	<b>35</b>
2.1 Intelligenz.....	35
2.2 Persönlichkeit.....	37
2.3 Motivation.....	39
2.4 Verhalten.....	43
2.5 Fazit.....	45
<b>3 Ausgewählte Ansätze zur Erklärung von Geschlechtsunterschieden in Motivation, Lernverhalten und Noten.....</b>	<b>47</b>
3.1 Erwartungs-Wert-Modell.....	48
3.1.1 Theoretische Annahmen.....	49
3.1.2 Bedeutung der Merkmale des Individuums.....	51
3.1.3 Bedeutung der sozialisierenden Umwelt.....	53
3.1.4 Zusammenfassung und Forschungsfazit.....	55
3.2 Identitätskongruente Nutzung schulischer Lernangebote.....	58
3.2.1 Theoretische Annahmen.....	58
3.2.2 Bedeutung einer Geschlechtsstereotypisierung von Schulfächern.....	61
3.2.3 Bedeutung einer Geschlechtsstereotypisierung von Schule und Lernen.....	63
3.2.4 Zusammenfassung und Forschungsfazit.....	67
3.3 Stereotypenbasierte Eindrucksbildung bei Lehrkräften.....	69
3.3.1 Wirkung von Stereotypen auf die Eindrucksbildung.....	69
3.3.2 Aktivierung von Geschlechterstereotypen durch Gender Enactment.....	72
3.3.3 Bedeutung des zugeschriebenen Verhaltens für Schulerfolg.....	74
3.3.4 Zusammenfassung und Forschungsfazit.....	76
<b>4 Fragestellungen, Ziele und Überblick der Studien.....</b>	<b>77</b>
<b>Literaturverzeichnis I.....</b>	<b>81</b>

---

<b>5 Studien .....</b>	<b>101</b>
5.1 Studie 1: Beyond expectancies and values: Applying the Eccles et al. model to the underachievement of boys .....	101
5.2 Studie 2: Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement .....	141
5.3 Studie 3: Do teachers equate male and masculine with lower academic engagement? How student's gender enactment triggers gender stereotypes at school.....	177
<b>6 Gesamtdiskussion .....</b>	<b>207</b>
6.1 Zusammenfassung der Studienergebnisse .....	207
6.2 Studienübergreifende Aspekte .....	210
6.2.1 Bedeutung von Geschlechterstereotypen .....	210
6.2.2 Bedeutung der geschlechtsbezogenen Identität.....	214
6.2.3 Bedeutung des untersuchten Fachs .....	216
6.2.4 Bedeutung des Schulerfolgsindikators Noten .....	218
6.3 Limitationen .....	219
6.4 Implikationen.....	222
6.4.1 Forschung .....	223
6.4.2 Praxis.....	228
6.5 Fazit.....	230
<b>Literaturverzeichnis II .....</b>	<b>233</b>
<b>Erklärung.....</b>	<b>241</b>
<b>Lebenslauf.....</b>	<b>243</b>

## Zusammenfassung

Aktuell besuchen Mädchen in Deutschland häufiger das Gymnasium als Jungen und erwerben höhere Schulabschlüsse, was zu einem großen Teil durch ihre besseren Deutsch- und Durchschnittsnoten erklärt werden kann. Mädchen verfügen außerdem über höhere Lesekompetenzen. Obwohl Jungen über leicht höhere mathematische Kompetenzen verfügen, erhalten sie nicht konsistent bessere Mathematiknoten. Bisherige Studien zur Erklärung der beschriebenen Geschlechtsunterschiede im Schulerfolg wiesen vor allem auf die Bedeutung des fachspezifischen Fähigkeitsselbstkonzepts, der fachspezifischen Werte sowie des fachübergreifend besseren Lernverhaltens und Benehmens der Mädchen als der Jungen hin. Im Rahmen dieser Arbeit wurden auf Grundlage des Erwartungs-Wert-Modells von Eccles et al. (z.B. Eccles et al., 1983), des Modells der Interessensentwicklung als Ausdruck der Identitätsregulation (z.B. Kessels & Hannover, 2004, 2006) sowie des Kontinuum-Modells der Eindrucksbildung (z.B. Fiske & Neuberg, 1990) weitere potentielle Erklärungen für den geringeren Schulerfolg der Jungen entwickelt und anschließend empirisch getestet.

*Studie 1* untersuchte an einer Stichprobe von  $N = 520$  Oberstufenschülerinnen und –schülern, ob bedeutende Variablen des Eccles et al.-Modells nicht nur Geschlechtsunterschiede in Mathematik zu Gunsten der Jungen, sondern auch in Deutsch zu Gunsten der Mädchen erklären können. Die zwei multiplen Mediationsmodelle mit dem Fähigkeitsselbstkonzept, den Werten, der Elterneinschätzung der kindlichen Fähigkeiten, der Intelligenz, dem Bildungsstand der Eltern und den früheren Noten als potentiell vermittelnde Variablen zeigten jedoch differenzielle Ergebnisse für Mathematik und Deutsch. Während die Geschlechtsunterschiede in den Mathematiknoten übererklärt wurden und Jungen unter Kontrolle der genannten Variablen nicht mehr bessere, sondern schlechtere Mathematiknoten als die Mädchen erhielten, wurden die besseren Deutschnoten der Mädchen nicht erklärt.

Angesichts der Bedeutung des selbstberichteten Lernverhaltens für die Erklärung von Geschlechtsunterschieden im Schulerfolg prüfte *Studie 2* an einer Stichprobe von  $N = 122$  Schülerinnen und Schülern der neunten Klasse, ob Schule und schulisches Lernen insgesamt feminin stereotypisiert wahrgenommen werden und in welchem Zusammenhang dies mit dem Schulerfolg von Jungen und Mädchen steht. Mithilfe eines computerbasierten impliziten Messverfahrens wurde gezeigt, dass Schule im Durchschnitt als leicht feminin wahrgenommen wurde. Diese Stereotypisierung ging bei den Jungen, die sich selbst stark mit sozial unerwünschten maskulinen Eigenschaften beschrieben, mit schlechteren Noten im ebenfalls feminin stereotypisierten Fach Deutsch einher. Mit den Mathematiknoten der Jungen oder mit den Mathematik- und Deutschnoten der Mädchen zeigten sich keine signifikanten Zusammenhänge. Dieses Ergebnis betont die Bedeutung der Passung von schulischen Angeboten und der geschlechtsbezogenen Identität für die Erklärung des geringeren Schulerfolgs der Jungen.

Da Lehrkräfte ihre Eindrücke bezüglich des Lernverhaltens von Schülerinnen und Schülern in ihre Notengebung miteinfließen lassen, wurde im Rahmen eines Vignettenexperiments mit  $N = 104$  Lehrkräften (*Studie 3*) untersucht, ob Gender Enactment von Jugendlichen stereotypenbasierte Erwartungen zu Ungunsten des Lernverhaltens und Benehmens von Jungen auslösen kann. Das Material war in drei Vorstudien entwickelt und vorgetestet worden ( $N = 104$ ,  $N = 82$ ,  $N = 86$ ). In Übereinstimmung mit den Hypothesen wurde Jungen insgesamt ein schlechteres Lernverhalten zugeschrieben als Mädchen. Gender Enactment verstärkte den Einfluss verhaltensbezogener Geschlechterstereotype auf die Eindrucksbildung. Jungen, die ihre Maskulinität betonten, wurde das schlechteste Lernverhalten zugeschrieben.

Die Ergebnisse der drei Studien weisen gemeinsam auf die Bedeutung (1) von fachlichen und verhaltensbezogenen Geschlechterstereotypen und (2) ihrer Interaktionen mit der geschlechtsbezogenen Identität und ihrer Entwicklung im Jugendalter für die Erklärung

von Geschlechtsunterschieden im Schulerfolg hin. Sie zeigen darüber hinaus auf, dass (3) Befunde aus einem Schulfach nicht ohne weiteres auf ein anderes Schulfach übertragbar sind und (4) sich auch hinter vermeintlich besseren Noten ein „relatives Underachievement“ im Sinne eines Nichtausschöpfens des eigenen Potentials verbergen kann. Abschließend werden weitere Fragestellungen für die zukünftige Forschung, die sich aus den Ergebnissen ableiten lassen, und Vorschläge für die schulische Praxis vorgestellt.



## Summary

In Germany, currently more girls than boys attend the Gymnasium, the academic-track school, and achieve the Abitur, the degree which qualifies for enrolling in a university. A large part of this gender effect can be explained by girls earning higher grades in German and higher GPAs. Girls also show higher reading competencies. Even though boys display slightly higher math competencies, they do not consistently earn higher math grades. Previous research explaining gender differences in academic achievement has stressed the role of the domain-specific ability self-concept, domain-specific subjective task values and girls' higher behavioral engagement compared to boys. Based on the expectancy-value model of Eccles et al. (e.g., Eccles et al., 1983), the interest as identity regulation model (e.g., Kessels & Hannover, 2007), and the continuum model of impression formation (e.g., Fiske & Neuberg, 1990), this dissertation derives and empirically tests additional hypotheses for explaining boys' lower academic achievement.

*Study 1* tested whether important variables of the Eccles et al. model could explain both boys' better math grades and girls' better German grades in a sample of  $N = 520$  German students attending the 11th and 12th grade. Students' ability self-concept and task values in math and German, intelligence and prior achievement, as well as parents' education and perception of the child's competence in math and German were included as multiple mediators in two structural equation models. Domain-specific results were found: After controlling for motivational, cognitive, and background variables, boys earned lower grades not only in German but also in math. Their lower grades in German remained unexplained by the mediators.

Given the importance of self-reported academic engagement for explaining gender differences in academic achievement, *Study 2* assessed whether school and academic engagement were stereotyped as feminine and how this gender stereotyping might relate to

boys' academic achievement in a sample of  $N = 122$  German 9th graders. Using a computer-based implicit measure it was found that, overall, students associated school and academic engagement slightly more with female than male. The more strongly boys associated school with female and the more they ascribed negative masculine traits to themselves, the lower their grades in the also feminine stereotyped subject German were. Boys' academic achievement in math was unrelated to the extent to which they perceived school as feminine, as were girls' grades in both German and math.

Teachers incorporate their impression regarding students' effort and engagement into their grading practice. Against this background *Study 3*, an experimental vignette study with  $N = 104$  academic track teachers, tested whether students' Gender Enactment triggered teachers' behavioral gender stereotypes resulting in an unfavorable perception of boys. Study materials were developed and pretested in three pilot studies ( $N = 104$ ,  $N = 82$ ,  $N = 86$ ). As hypothesized, teachers generally ascribed less behavior fostering learning and more behavior impeding learning to boys than to girls. Gender Enactment, however, amplified the impact of the behavioral gender stereotypes on impression formation. Boys enacting masculinity were rated as showing the most behavior impeding learning and the least behavior fostering learning.

Taken together, the results point to the importance of (1) domain-specific and behavioral gender stereotypes, and (2) their interaction with students' (developing) gender identity in adolescence for explaining gender differences in academic achievement. Further, they indicate that (3) results found in one academic domain cannot simply be generalized to another domain and (4) even on first glance better grades can conceal an academic underachievement relative to a student's academic potential. Finally, based on the findings, directions for future research and tentative suggestions for the classroom are proposed.



## Einleitung

„Sind Jungen die neuen Bildungsverlierer?“ Unter dieser Fragestellung werden mit großer medialer Aufmerksamkeit seit mehr als einem Jahrzehnt in Deutschland Geschlechtsunterschiede im Schulerfolg zu Ungunsten von Jungen diskutiert. Da Jungen und Mädchen in der Schule jedoch nicht wie zwei Sportmannschaften im Kampf um Bildung gegeneinander antreten, ist offensichtlich, dass die so formulierte Frage im besten Falle provozieren und zur tieferen Auseinandersetzung mit dem Phänomen anregen kann. Eine einfache Beantwortung dieser Frage mit „ja“ oder „nein“ kann weder die Komplexität des Konzepts Schulerfolg, noch die des Konzepts Geschlecht, noch die differenziellen Zusammenhänge zwischen beiden abbilden. Vor allem aber wird sie den Kindern und Jugendlichen, die zurzeit in Deutschland zur Schule gehen, und ihren Lehrkräften nicht gerecht.

Diese Arbeit hat zum Ziel, der pauschalisierenden Darstellung von Jungen als „Bildungsverlierern“ eine differenzierte, empirisch fundierte Analyse gegenwärtiger Geschlechtsunterschiede im Schulerfolg entgegenzusetzen, sowie mithilfe psychologischer Theorien und Methoden einen wissenschaftlichen Beitrag zu ihrer Erklärung zu leisten. Denn hinter der Frage nach „Bildungsverlierern“ verbirgt sich trotz ihrer simplifizierten Formulierung die bedeutende öffentliche und wissenschaftliche Debatte um Bildungsgerechtigkeit in Deutschland. Bildungsgerechtigkeit ist ein großes Ideal und ein wesentliches Ziel der deutschen Bildungspolitik (BMBF, 2015). Unterschiede im Schulerfolg, die von der ethnischen, kulturellen oder sozialen Herkunft, Muttersprache, Religion oder vom Geschlecht abhängig sind statt von den kognitiven Leistungen, stellen die Bildungsgerechtigkeit in Deutschland in Frage. Diese Arbeit fokussiert eines dieser soziodemografischen Merkmale, das Geschlecht, um zu einem fundierteren Verständnis der dem differenziellen Schulerfolg von Jungen und Mädchen zu Grunde liegenden Prozesse

beizutragen. Denn nur aus solch einem Verständnis heraus können effektive Handlungsempfehlungen zur Steigerung der Bildungsgerechtigkeit in Deutschland abgeleitet werden.

Der Fokus dieser Arbeit liegt auf der Rolle, die individuelle psychologische Merkmale und gesellschaftlich geteilte Geschlechterstereotype für die Erklärung des differenziellen Schulerfolgs von Mädchen und Jungen spielen. Dabei werden verschiedene Perspektiven auf „Geschlecht“ berücksichtigt und zusammengebracht. Geschlecht wird sowohl als Personenmerkmal mit den Ausprägungen männlich und weiblich („gender as a subject variable perspective“), als psychologisches Merkmal mit den Dimensionen Maskulinität und Femininität, die mit männlichen und weiblichen Personen assoziiert sind („gender as a personality variable perspective“), als auch als soziale Kategorie betrachtet („gender as a social category perspective“; Eckes & Trautner, 2000a). Letztere Perspektive ermöglicht es, nicht nur das handelnde Individuum, sondern auch die Bedeutung der Beobachtenden und der Situation bei der Entstehung von Geschlechtsunterschieden zu berücksichtigen. Keine der drei Perspektiven impliziert, dass Unterschiede zwischen Mädchen und Jungen sowie Annahmen darüber unveränderlich oder „natürlich“ seien. Biologische, z.B. auf Hormonen basierende Erklärungen für Unterschiede zwischen Mädchen und Jungen sind nicht Bestandteil der Arbeit, ohne dass ihr Einfluss auf menschliches Verhalten damit negiert werden soll.

Die Arbeit besteht aus drei empirischen Studien, denen eine gemeinsame theoretische Einleitung vorangestellt ist und an die eine gemeinsame Diskussion der Studienergebnisse anschließt. Im Kapitel 1 des Theorieteils werden zunächst die gegenwärtigen Unterschiede zwischen dem Schulerfolg der Mädchen und dem der Jungen am Beispiel der Indikatoren (1) Bildungsbeteiligung und erworbene Schulabschlüsse, (2) Schulnoten und (3) standardisierte Leistungs- und Kompetenztests vorgestellt. Im nachfolgenden Kapitel 2 wird zusammengefasst, inwieweit sich die beschriebenen Geschlechtsunterschiede im Schulerfolg durch mögliche Geschlechtsunterschiede in individuellen, psychologischen,

schulerfolgsrelevanten Merkmalen erklären lassen. In Kapitel 3 werden auf Grundlage des Erwartungs-Wert-Modells von Eccles et al. (z.B. Eccles et al., 1983), des Modells der Interessensentwicklung als Ausdruck der Identitätsregulation (z.B. Kessels & Hannover, 2004, 2006) sowie des Kontinuum-Modells der Eindrucksbildung (Fiske & Neuberg, 1990) weitere potentielle Erklärungen für den geringeren Schulerfolg der Jungen entwickelt. Kapitel 4 fasst die sich daraus ableitenden Fragestellungen zusammen und gibt einen kurzen Überblick über die Ziele und Methodik der drei empirischen Studien, die zu ihrer Prüfung konzipiert und durchgeführt wurden. Die drei Studien bilden, jeweils in Manuskriptform, das Kapitel 5. An die Manuskripte schließt sich mit Kapitel 6 eine gemeinsame Diskussion an, die zunächst zentrale Ergebnisse der einzelnen Studien zusammenfasst und anschließend studienübergreifende Ergebnisse und Limitationen vor dem Hintergrund der bisherigen Forschung diskutiert. Es werden weitere Forschungsperspektiven eröffnet und mögliche Vorschläge für die schulische Praxis entwickelt. Die Gesamtdiskussion endet mit einem studienübergreifenden Fazit.



## **1 Geschlechtsunterschiede im Schulerfolg**

Im Gegensatz zur vereinfachenden medialen Darstellung von Jungen als „Verlierern“ und Mädchen als „Gewinnern“ in der Schule unterscheiden sich die Geschlechtsunterschiede im Schulerfolg in Abhängigkeit davon, welcher Indikator für Schulerfolg betrachtet wird.

Basierend auf dem Überblicksartikel von Hannover und Kessels (2011) werden im Folgenden die Bildungsbeteiligung und die erworbenen Schulabschlüsse von Mädchen und Jungen in Prozentwerten sowie ihre durchschnittlichen Noten und fachlichen Kompetenzen als Indikatoren für Schulerfolg herangezogen und verglichen. Auch wenn Geschlechtsunterschiede in der Variabilität der Indikatoren für Schulerfolg nicht behandelt werden, ist bei der Interpretation selbstverständlich zu berücksichtigen, dass Unterschiede zwischen Mittelwerten in der berichteten Größe deutlich überlappende Merkmalsverteilungen zwischen Mädchen und Jungen implizieren (z.B. Hyde, 2005). Trotz Mittelwertunterschieden schneiden viele Mädchen und Jungen also gleich gut ab. Ebenso erreichen viele Mädchen und Jungen gleichwertige Schulabschlüsse.

### **1.1 Bildungsbeteiligung und Schulabschlüsse**

Seit den 1950er Jahren steigt in Deutschland kontinuierlich der Anteil der Jugendlichen, die das Gymnasium besuchen und die allgemeine Hochschulreife erwerben. Im gleichen Zeitraum wuchs der Anteil der Mädchen, die ein Gymnasium besuchen und das Abitur erwerben, stärker als der Anteil der Jungen, die ein Gymnasium besuchen und das Abitur erwerben, sodass ungefähr seit 1980 Mädchen am Gymnasium überrepräsentiert sind und höherwertige Schulabschlüsse erwerben (Statistisches Bundesamt, 1982, 1983). Im Folgenden werden die Statistiken des Statistischen Bundesamts zur Bildungsbeteiligung (d.h. Schulformbesuch) und zu den erworbenen Schulabschlüssen von Jungen und Mädchen beispielhaft für das Schuljahr 2012/2013 zusammengefasst. In den Schuljahren davor zeigte sich ein ähnliches Bild. Laut des Statistischen Bundesamts (2013a) besuchten im Schuljahr

2012/2013 ca. 8,6 Millionen Kinder und Jugendliche in Deutschland allgemeinbildende Schulen. Davon waren 51% Jungen. Jedoch verteilten sich Jungen und Mädchen nicht entsprechend dieses Verhältnisses auf die verschiedenen Schulformen. Mädchen besuchten überdurchschnittlich häufig das Gymnasium. Der Mädchenanteil an der gymnasialen Schülerschaft betrug 52,5%, der Jungenanteil entsprechend 47,5%. Jungen stellten am Gymnasium folglich eine etwas kleinere, stärker selektierte Gruppe dar als Mädchen. Die anderen Schulformen im Sekundarbereich 1 besuchten hingegen mehr Jungen als Mädchen (mit Ausnahme der Freien Waldorfschule). Besonders deutlich war der Unterschied bei den Haupt- und Förderschulen (Hauptschule: Jungenanteil 56,4 %; Förderschule: Jungenanteil 64,0%). Außerdem wiederholte in jeder Schulform ein größerer Anteil der Jungen als der Mädchen ein Schuljahr (insgesamt 2,7% der Jungen im Vergleich zu insgesamt 1,9% der Mädchen).

Auch bei den erworbenen Schulabschlüssen lagen Mädchen vorn. Im Schuljahr 2012/2013 waren 54,7% derjenigen, die das Abitur erworben haben, weiblich und entsprechend 45,3% männlich (Statistisches Bundesamt, 2013a). Unter denjenigen, die einen Hauptschulabschluss erworben oder die Schule ohne Abschluss verlassen haben, betrug der Jungenanteil hingegen 58,2% bzw. 60,3%. Die auf den ersten Blick möglicherweise gering erscheinenden prozentualen Unterschiede werden bedeutender, wenn zusätzlich die absoluten Zahlen berücksichtigt werden. So erlangten im Schuljahr 2012/2013 z.B. 28.634 mehr Mädchen als Jungen das Abitur, während 9.814 mehr Jungen als Mädchen die Schule ohne Abschluss verließen (Statistisches Bundesamt, 2013a).

## **1.2 Schulnoten**

Schulnoten stellen das zentrale Erfolgs- und Selektionskriterium im Schulkontext dar (z.B. Tent & Birkel, 2010; Ziegenspeck, 1999). Sie beeinflussen, welche weiterführende Schule nach der Grundschule besucht wird, ob ein Schuljahr wiederholt werden soll oder welcher

Abschluss erreicht wird (z.B. für Berlin: Verordnung über die Schularten und Bildungsgänge der Sekundarstufe I, 2010). Schulgesetze bilden die rechtliche Grundlage der Notenvergabe. Im Bundesland Nordrhein-Westfalen<sup>1</sup> sollen z.B. Noten „über den Stand des Lernprozesses der Schülerin oder des Schülers Aufschluss geben“ (Schulgesetz für das Land Nordrhein-Westfalen, 2005, §48 (1)) und „sich auf die im Unterricht vermittelten Kenntnisse, Fähigkeiten und Fertigkeiten“ beziehen (Schulgesetz für das Land Nordrhein-Westfalen, 2005, §48 (2)). Sie basieren auf den von der Schülerin oder dem Schüler in einem bestimmten Zeitraum schriftlich und mündlich erbrachten Leistungen (Schulgesetz für das Land Nordrhein-Westfalen, 2005, §48 (2)).

Internationale Befragungen von Lehrkräften und experimentelle Vignettenstudien zeigen jedoch, dass Lehrkräfte neben den gezeigten Leistungen der Schülerinnen und Schüler auch ihre oder seine (vermutete) Begabung, gezeigte Anstrengung oder Verhalten im Unterricht in die Notengebung miteinfließen lassen (Brookhart, 1993, 1994; Cizek, Fitzgerald & Rachor, 1995; Duckworth, Quinn & Tsukayama, 2012; McMillan, 2001; McMillan, Myran & Workman, 2002; Randall & Engelhard, 2010; Sun & Cheng, 2014). Die Lehrkräfte gehen bei der Notengebung jedoch sehr durchdacht vor und bemühen sich um Fairness und positive Konsequenzen für die Motivation und den Selbstwert der Schülerin oder des Schülers (Brookhart, 1993; Sun & Cheng, 2014). So kommen sie der Motivations-, Disziplinar- und Rückmeldefunktion von Noten nach (z.B. Tent & Birkel, 2010). Letztere spiegelt sich auch in den mittelstarken bis starken Zusammenhängen zwischen Noten und selbsteingeschätzten Fähigkeiten der Schülerinnen und Schüler wider (z.B. Möller, Pohlmann, Köller & Marsh, 2009). Neben weiteren Merkmalen der Schülerin oder des Schülers fließt auch das Leistungsniveau der Mitschülerinnen und Mitschüler in die Notengebung ein, was die absolute Vergleichbarkeit von Noten zwischen Klassen verringert (z.B. Ingenkamp & Lissmann, 2008). Diese (pädagogisch teilweise erwünschte) Berücksichtigung von Faktoren

---

<sup>1</sup> Nordrhein-Westfalen wird an dieser Stelle beispielhaft vorgestellt, da die Daten der empirischen Studien dieser Arbeit größtenteils in diesem Bundesland erhoben wurden.

über Leistung hinaus führt zu einer geringeren Objektivität und Reliabilität von Noten, als z.B. standardisierte Kompetenztests sie aufweisen (z.B. Ingenkamp & Lissmann, 2008; Willingham, Pollack & Lewis, 2002). Gleichzeitig besitzen Noten eine höhere ökologische Validität als standardisierte Leistungstests (Schilling, Sparfeldt & Rost, 2006; Spinath, Eckert & Steinmayr, 2014). Die mittelstarken Zusammenhänge zwischen Lehrerbeurteilungen, z.B. in Form von Noten, und Schülerleistungen in standardisierten Leistungstests im gleichen Fach (Hülür, Wilhelm & Robitzsch, 2011; Lüdtke, Köller, Artelt, Stanat & Baumert, 2002; Südkamp, Kaiser & Möller, 2012) sowie die Bedeutung von Noten für Studien- und Berufserfolg (Bowen, Chingos & McPherson, 2011; Schuler, 2010; Trapmann, Hell, Weigand & Schuler, 2007) untermauern ihre Konstrukt- und prognostische Validität.

Bereits Hannover und Kessels (2011) diskutierten, ob die geringere Bildungsbeteiligung der Jungen in Deutschland darauf zurückzuführen sei, dass Jungen schlechtere Noten erhielten als Mädchen und somit seltener eine Gymnasialempfehlung ausgesprochen bekämen. Im Gegensatz zu Schulabschlüssen sind Schulnoten nach Kenntnis der Autorin der vorliegenden Arbeit in keiner bundesweiten, öffentlich verfügbaren Statistik erfasst, sodass zur Klärung der Frage nach Geschlechtsunterschieden in Schulnoten auf einzelne Studien mit deutschen Stichproben zurückgegriffen werden muss.

Hinsichtlich etwaiger Geschlechtsunterschiede in der Mathematiknote berichteten verschiedene Studien, die auf großen, teilweise repräsentativen deutschen Stichproben basieren, gemischte Befunde. Am Ende der Grundschulzeit erhielten Jungen im Durchschnitt etwas bessere Mathematiknoten als Mädchen (z.B. ÜBERGANG-Studie des Max-Planck-Instituts Berlin Maaz, Baeriswyl & Trautwein, 2011; PIRLS 2001 Neugebauer, 2011). Dieser Unterschied erreichte jedoch nicht in allen Fällen statistische Signifikanz (z.B. KEGS-Studie Kuhl & Hannover, 2012; ELEMENT-Studie Helbig, 2010b). Für den Sekundarschulbereich werden die Ergebnisse etwas heterogener. Die meisten Studien berichteten von besseren Mathematiknoten der Jungen als der Mädchen mit kleinen Effektgrößen (Dresel, Stoeger &



Ziegler, 2006; Dumont et al., 2012; Steinmayr, Wirthwein & Schöne, 2014). Jedoch fanden sich teilweise auch keine (signifikanten) Geschlechtsunterschiede (Köller, Daniels, Schnabel & Baumert, 2000; Schilling et al., 2006; Steinmayr & Spinath, 2008) oder sogar bessere Mathematiknoten der Mädchen als der Jungen (Freund, Holling & Preckel, 2007; Steffens & Jelenec, 2011). Dies spiegelt sich auch in der aktuellen Metaanalyse von Voyer und Voyer (2014) zu Geschlechtsunterschieden in Schulnoten wider, die Mädchen in nordamerikanischen Studien sogar bessere Mathematiknoten attestiert als Jungen. Für nicht-nordamerikanische Studien wurden keine signifikanten Geschlechtsunterschiede festgestellt (Voyer & Voyer, 2014). Leider wurden hier die Studienländer nur in die drei Kategorien „Nordamerika“, „Skandinavien“ und „Sonstige“ aufgeschlüsselt, sodass die Metaanalyse keine spezifischen Aussagen zu Deutschland trifft.

Im Fach Deutsch zeigte sich hingegen ein eindeutiges Bild: Hier erhielten Mädchen konsistent und unabhängig von der betrachteten Jahrgangsstufe bessere Noten als Jungen (Dresel & Ziegler, 2006; Freund et al., 2007; Helbig, 2010b; Kuhl & Hannover, 2012; Maaz et al., 2011; Neugebauer, 2011; Schilling et al., 2006; Steffens & Jelenec, 2011; Steinmayr & Spinath, 2008). In der internationalen Metaanalyse wurden ebenfalls signifikante Unterschiede in den Noten des muttersprachlichen Fachs zu Gunsten der Mädchen gefunden, die für nordamerikanische Studien und die Sekundarstufe noch einmal größer ausfielen als für Studien aus anderen Ländern und den Primarbereich (Voyer & Voyer, 2014).

Darüber hinaus erhielten Schülerinnen international und in Deutschland bessere Durchschnittsnoten als Schüler (Voyer & Voyer, 2014; für Deutschland 4. Klasse z.B. Maaz et al., 2011; für Abiturdurchschnitt: z.B. Ministerium für Schule und Weiterbildung NRW, 2013). Im Gegensatz zur Bildungsbeteiligung der Frauen und Mädchen (siehe Kapitel 1.1) sind ihre Durchschnittsnoten in Deutschland in den vergangenen Jahrzehnten jedoch nicht bedeutend gestiegen: Schon in den 50er und 60er Jahren des vergangenen Jahrhunderts haben

Frauen und Mädchen im Durchschnitt bessere Schulnoten erhalten als Männer und Jungen (Rodax & Hurrelmann, 1986; Ziegenspeck, 1999, S. 145).

Die besseren Deutsch- und Durchschnittsnoten der Mädchen tragen dazu bei, dass die Chance eines Mädchens, eine Empfehlung für das Gymnasium zu erhalten, größer ist als die eines Jungen (Hannover & Kessels, 2011; Neugebauer, 2011). Auch Ditton (2007) fand nach Kontrolle der Noten keine Geschlechtsunterschiede mehr in der Grundschulempfehlung. Allerdings waren die zu erklärenden Geschlechtsunterschiede in den Grundschulempfehlungen in der untersuchten Stichprobe von Anfang an klein. Im Fächervergleich scheint der Deutschnote bei der Vorhersage einer Empfehlung für das Gymnasium mehr Gewicht zuzukommen als der Mathematiknote (z.B. KEGS-Studie Fuchs, 2014; ELEMENT-Studie Helbig, 2010b; LAU-Studie Lehmann, Peek & Gänsfuß, 1997; TIMSS-Übergang Maaz & Nagy, 2010; für die Schweiz und Befunde für den Übergang in die Sekundarstufe 2 siehe z.B. Neuenschwander & Malti, 2009). Dies trägt auch zur höheren Präsenz der Mädchen an deutschen Gymnasien bei (Neugebauer, 2011). Den besseren Deutsch- und Durchschnittsnoten der Mädchen am Ende der Grundschulzeit kommt folglich eine zentrale Rolle bei der Erklärung von Geschlechterdisparitäten in Bildungsbeteiligung und Schulabschlüssen zu (Hannover & Kessels, 2011).

### **1.3 Leistungs- und Kompetenztests**

Die Ergebnisse standardisierter Leistungs- und Kompetenztests stellen einen weiteren Indikator für Schulerfolg dar (Hannover & Kessels, 2011). Diese werden in regelmäßigen Abständen an repräsentativen Stichproben im Rahmen nationaler und internationaler Schulleistungsuntersuchungen erfasst. Wegen der Repräsentativität der Stichprobe und der Standardisierung des Tests können Aussagen über den Leistungsstand aller deutschen Schülerinnen und Schüler gemacht werden und mit denen in anderen Ländern, zwischen Schultypen, Schülergruppen u.v.m. verglichen werden. Im Gegensatz zu Zeugnisnoten,

welche auf einem breiten Leistungsverständnis basieren und die innerhalb eines Halbjahres gezeigten Leistungen abbilden, messen Schulleistungsuntersuchungen die Kompetenzen und Leistungen mit spezifischen, standardisierten Tests in einer Domäne zu einem Testzeitpunkt (z.B. Willingham et al., 2002). Das Abschneiden in standardisierten Leistungstests ist stärker von der Intelligenz der Schülerin oder des Schülers und weniger stark von anderen überfachlichen Kompetenzen, wie z.B. schulischem Engagement, Selbstkontrolle oder angemessenem Verhalten im Unterricht, beeinflusst als Noten (Duckworth et al., 2012; Hofer, Kuhnle, Kilian & Fries, 2012; McDermott, Mordell & Stoltzfus, 2001; Steinmayr & Meißner, 2013; Willingham et al., 2002). Somit messen Noten und Leistungs- oder Kompetenztests leicht verschiedene Aspekte von Schulerfolg, haben jedoch einen gemeinsamen Kern (Willingham et al., 2002), der sich auch in den mittelstarken, positiven Korrelationen zwischen beiden Indikatoren zeigt (z.B. Hülür et al., 2011; Lüdtke et al., 2002; Willingham et al., 2002).

Im Folgenden werden die Kompetenz- und Leistungsunterschiede zwischen Jungen und Mädchen in der Effektgröße Cohen's  $d$  dargestellt. In Anlehnung an Cohen (1988) werden häufig Unterschiede um  $d = .2$  als klein, um  $d = .5$  als mittelgroß und um  $d = .8$  als groß interpretiert. Um die praktische Relevanz eines Unterschieds abzuschätzen, ist jedoch der Vergleich mit relevanten Benchmarks wie dem zu erwartenden Lernzuwachs in einem Schuljahr viel nützlicher (z.B. Bloom, Hill, Black & Lipsey, 2008; Hill, Bloom, Black & Lipsey, 2008). Zu diesem Zweck wird zusätzlich die Größe des Lernzuwachses innerhalb eines Schuljahres berichtet, die im Rahmen der Pilotierung und Normierung der Bildungsstandards für das vierte Schuljahr und im Rahmen des PISA-Längsschnitts 2003/2004 für 15-Jährige festgestellt wurde. Im Folgenden werden zunächst Ergebnisse aus der Primarstufe und anschließend aus der Sekundarstufe vorgestellt.

Die Schulleistungsuntersuchung Trends in International Mathematics and Science Study (TIMSS) wird von der International Association for the Evaluation of Educational

Achievement (IEA) verantwortet und untersucht im Vier-Jahres-Rhythmus die mathematischen und naturwissenschaftlichen Kompetenzen von Schülerinnen und Schülern in der vierten und achten Jahrgangsstufe sowie am Ende der Sekundarstufe 2. Dabei orientieren sich die Testaufgaben stark an den nationalen Lehrplänen, um im Sinne der intendierten Curricula valide zu sein (z.B. Köller, Baumert & Bos, 2002). Deutschland beteiligte sich beim letzten Durchgang im Jahr 2011 (wie auch bei früheren Messungen) nur mit der vierten Jahrgangsstufe, sodass TIMSS 2011 nur Informationen zu den Kompetenzen deutscher Schülerinnen und Schüler in Mathematik und den Naturwissenschaften am Ende des Primarbereichs liefern kann. Die TIMSS-Ergebnisse von 2011 zeigten, dass Jungen in der vierten Klasse in Deutschland im Bereich Mathematik einen Kompetenzvorsprung in der Größe von  $d = .13$  besaßen (eigene Berechnung [im Folgenden: e. B.] auf Grundlage von Selter, Walther, Wessel & Wendt, 2012 und Brehl, Wendt & Bos, 2012). Die Unterschiede variierten jedoch zwischen den verschiedenen mathematischen Inhaltsbereichen und kognitiven Anforderungsbereichen (Brehl et al., 2012). Im Durchschnitt der Mitgliedsländer der OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung) betrug der Unterschied  $d = .08$  zu Gunsten der Jungen (e. B. auf Grundlage von Selter et al., 2012 und Brehl et al., 2012). In etwa der Hälfte der untersuchten Länder lagen keine signifikanten Geschlechtsunterschiede in der Mathematikleistung vor (Brehl et al., 2012). Beim IQB-Ländervergleich der Bildungsstandards am Ende der Primarstufe fielen die Geschlechtsunterschiede in der globalen mathematischen Kompetenz mit  $d = .16$  zu Gunsten der Jungen (Böhme & Roppelt, 2012) ähnlich klein aus wie in der deutschen TIMSS-Stichprobe (für Unterschiede und Gemeinsamkeiten der Studien siehe Böhme et al., 2014). Ein ähnliches Bild zeigte sich für den naturwissenschaftlichen Bereich. Auch hier lagen bei TIMSS 2011 in etwa der Hälfte der Länder keine signifikanten Geschlechtsunterschiede vor (Brehl et al., 2012). Im OECD-Durchschnitt betrug der Unterschied  $d = .08$  zu Gunsten der Jungen (e. B. auf Grundlage von Kleickmann, Brehl, Saß, Prenzel & Köller, 2012 und Brehl

et al., 2012). In Deutschland schnitten Jungen im Durchschnitt auch etwas besser ab als Mädchen ( $d = .17$ , e. B. auf Grundlage von Kleickmann et al., 2012 und Brehl et al., 2012). Die Unterschiede zwischen den mathematischen und naturwissenschaftlichen Kompetenzen der Jungen und denen der Mädchen in der vierten Klasse sind als sehr klein zu bewerten (Cohen, 1988). Als weitere interpretative Heuristik kann die Leistungsdifferenz zwischen den Schülerinnen und Schülern des dritten und denen des vierten Schuljahrs aus den Studien zur Pilotierung und Normierung der Bildungsstandards im Primarbereich genutzt werden (Selter et al., 2012). Diese betrug  $d = .84$  im Fach Mathematik (Reiss & Winkelmann, 2009), welches die Interpretation des Geschlechtsunterschieds in TIMSS in der 4. Klasse als „klein“ stärkt. In amerikanischen Tests und Stichproben fanden sich für die gleiche Jahrgangsstufe tendenziell etwas kleinere Leistungszuwächse in der Höhe von etwa einer halben Standardabweichung (Bloom et al., 2008; Hill et al., 2008).

Die Progress in International Reading Literacy Study (PIRLS) wird ebenfalls von der IEA durchgeführt und erfasst in den gleichen Jahrgangsstufen wie TIMSS alle fünf Jahre das Leseverständnis. Wie auch bei TIMSS beteiligte sich Deutschland im Jahr 2011 nur mit der vierten Jahrgangsstufe. Die PIRLS-Ergebnisse von 2011 zeigten, dass im Großteil der untersuchten Länder Mädchen am Ende der vierten Klasse signifikant bessere Leseleistungen aufwiesen als Jungen (Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Im OECD-Durchschnitt betrug der Kompetenzvorsprung der Mädchen im Lesen in der vierten Klasse  $d = .17$  (e. B. auf Grundlage von Bos et al., 2012). Nur in vier Ländern zeigten sich keine signifikanten Unterschiede. Allerdings variierten die Geschlechtsunterschiede in ihrer Größe deutlich zwischen den Ländern. Während die Mädchen in Finnland beispielsweise einen Vorsprung von  $d = .33$  besaßen, fielen die Geschlechtsunterschiede in Deutschland mit  $d = .12$  zu Gunsten der Mädchen vergleichsweise klein aus (e. B. auf Grundlage von Bos et al., 2012). Im IQB-Ländervergleich 2011 zeigten sich mit  $d = .24$  etwas größere Unterschiede im Lesen zu Gunsten der Mädchen (Böhme & Roppelt, 2012; für Unterschiede und Gemeinsamkeiten

der Studien siehe Böhme et al., 2014). Als Interpretationshilfe kann vorsichtig die bei der Pilotierung und Normierung der Bildungsstandards ermittelte Leistungsdifferenz im Lesen zwischen der dritten und der vierten Jahrgangsstufe von  $d = .60$  herangezogen werden (Bremerich-Vos & Böhme, 2009).

Insgesamt lässt sich für deutsche Schülerinnen und Schüler im vierten Schuljahr folglich festhalten, dass sich ihre Kompetenzen nur in geringem Ausmaß unterscheiden: Jungen verfügen über geringfügig höhere mathematische und naturwissenschaftliche Kompetenzen, Mädchen über geringfügig höhere Lesekompetenzen. Auch wenn die Geschlechtsunterschiede im Lesen zu Gunsten der Mädchen in Deutschland, wenn sie in Cohen's  $d$  dargestellt werden, ähnlich groß ausfallen wie die in der mathematischen Kompetenz zu Gunsten der Jungen, können sie tentativ als etwas bedeutender interpretiert werden, da sowohl die Studien zur Pilotierung und Normierung der Bildungsstandards im Primarbereich (Bremerich-Vos & Böhme, 2009) als auch die Ergebnisse amerikanischer Leistungstests (Hill et al., 2008) darauf hindeuten, dass im Lesen kleinere Leistungszuwächse zwischen der dritten und vierten Klasse zu erwarten sind als in der Mathematik.

Die größte Schulleistungsuntersuchung im Sekundarschulbereich, die Studie Programme for International Student Assessment (PISA), testet alle drei Jahre in einer stetig steigenden Anzahl von Ländern die Lesekompetenz, die mathematische Kompetenz und die naturwissenschaftliche Kompetenz der 15-Jährigen. Weitere Tests stehen optional zur Verfügung (für eine ausführlichere Darstellung siehe z.B. Sälzer & Prenzel, 2013). PISA wird im Auftrag der OECD durchgeführt. Statt der Überprüfung schulischer Curricula versucht PISA abzubilden, inwieweit 15-Jährige (in der Schule) erworbene Schlüsselkompetenzen auf Problemstellungen anwenden können, um erfolgreich am gesellschaftlichen Leben teilzuhaben (z.B. Sälzer & Prenzel, 2013). Damit unterscheidet sich die PISA-Studie in ihrer Anlage und ihrem Ziel grundlegend von anderen Schulleistungsuntersuchungen wie z.B. TIMSS. Außerdem basiert PISA auf einer Alters- statt einer Jahrgangsstufenstichprobe,

sodass die getesteten 15-Jährigen die siebte bis elfte Klasse besuchen können. Eine auf der PISA-Messwiederholung von 2003/2004 basierende Faustregel besagt, dass Unterschiede in der Größe von  $d = .25$  bis  $d = .30$  ungefähr dem durchschnittlichen Lernzuwachs innerhalb eines Schuljahres entsprechen (Prenzel et al., 2006; Sälzer & Prenzel, 2013).

In jeder PISA-Erhebung wird einer der Kompetenzbereiche, die sogenannte Hauptdomäne, detaillierter getestet. Im Jahr 2012 lag der Schwerpunkt auf Mathematik. Hier betrug der Kompetenzunterschied zwischen Jungen und Mädchen im OECD-Durchschnitt  $d = .11$  zu Gunsten der Jungen (e. B. auf Grundlage von Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013). Die Größe und Richtung dieses Geschlechtsunterschieds entsprach den Befunden internationaler Metaanalysen (z.B. Else-Quest, Hyde & Linn, 2010; Reilly, Neumann & Andrews, 2014). Ebenso wie die genannten Metaanalysen zeigten die Ergebnisse von PISA 2012 außerdem deutlich, dass die Geschlechtsunterschiede in ihrer Größe zwischen den Ländern variierten. In ca. einem Drittel der teilnehmenden Länder zeigten sich bei PISA 2012 keine signifikanten Unterschiede in den Mathematikkompetenzen zwischen Jungen und Mädchen, vereinzelt erreichten Mädchen auch höhere Kompetenzwerte als Jungen (Sälzer et al., 2013). In Deutschland schnitten Jungen in den mathematischen Kompetenztests etwas besser ab als Mädchen ( $d = .14$ , e. B. auf Grundlage von Sälzer et al., 2013). Da Mädchen an höheren Schulformen überrepräsentiert sind (siehe Kapitel 1.1) und somit häufiger einer besonders fördernden und anspruchsvollen Lernumgebung ausgesetzt sind (Baumert, Stanat & Watermann, 2006b), vergrößerten sich die Geschlechtsunterschiede in der Mathematikkompetenz zu Gunsten der Jungen, wenn die Schulform berücksichtigt wurde (OECD, 2013a, S. 340). Ähnliches war bereits für PISA 2000 (Stanat & Kunter, 2001) und TIMSS 1995 berichtet worden (Hosenfeld, Köller & Baumert, 1999; siehe auch Becker, Lüdtke, Trautwein & Baumert, 2006). Wurden die Werte der Mathematiktests früherer PISA-Durchgänge für die allgemeine Intelligenz kontrolliert, zeigten sich außerdem größere Geschlechtsunterschiede zu Gunsten der Jungen in der *spezifischen* mathematischen

Kompetenz als in der nicht für die Intelligenz kontrollierten mathematischen Kompetenz (Brunner et al., 2013; Brunner, Krauss & Kunter, 2008).

Wie bei PIRLS 2011 schnitten Mädchen auch in PISA 2012 im Lesen im OECD-Durchschnitt besser ab als Jungen ( $d = .39$ ; e. B. auf Grundlage von Hohn, Schiepe-Tiska, Sälzer & Artelt, 2013). Jedoch fiel der Unterschied bei den 15-Jährigen im OECD-Durchschnitt deutlich größer aus als der zwischen den Schülerinnen und Schülern der vierten Klasse (siehe oben). Außerdem erreichten die Mädchen in *allen* teilnehmenden Staaten mehr Lesekompetenzpunkte als Jungen. In Deutschland betrug der Unterschied 2012  $d = .48$  zu Gunsten der Mädchen (e. B. auf Grundlage von Hohn et al., 2013). Da Mädchen aufgrund ihrer Überrepräsentation an Gymnasien vermehrt einer besonders günstigen Lernumgebung ausgesetzt sind, verringerten sich z.B. bei PISA 2000 oder PISA 2009 die Geschlechtsunterschiede in der Lesekompetenz zu Gunsten der Mädchen, wenn die besuchte Schulform berücksichtigt wurde (OECD, 2010; Stanat & Kunter, 2001). Ein ähnliches Bild zeigte sich auch beim Ländervergleich zur Überprüfung der Bildungsstandards für den Mittleren Schulabschluss für das Fach Deutsch (Köller, Knigge & Tesch). Die naturwissenschaftlichen Kompetenzen der 15-jährigen Mädchen und Jungen unterschieden sich weder im OECD-Durchschnitt ( $d = .00$ ) noch in Deutschland ( $d = .02$  zu Gunsten der Jungen; e. B. auf Grundlage von Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013).

Insgesamt zeigten sich bei den deutschen 15-Jährigen die größten – aber immer noch als klein bis mittelstark einzuschätzenden – Geschlechtsunterschiede in der Lesekompetenz zu Gunsten der Mädchen, kleine Unterschiede in der mathematischen Kompetenz zu Gunsten der Jungen und keine Unterschiede in der naturwissenschaftlichen Kompetenz. Tendenziell fanden sich – ähnlich wie bei den Schulnoten (Voyer & Voyer, 2014) – größere Geschlechtsunterschiede im Sekundar- als im Primarbereich im Sinne eines Schereneffekts (Dresel et al., 2006; Hill et al., 2008). Aufgrund des Querschnittsdesigns von PISA, TIMSS und dem IQB-Ländervergleich und den Unterschieden in der Studienkonzeption (Böhme et



al., 2014; van den Ham, Nissen, Ehmke, Sälzer & Roppelt, 2014), ist diese Aussage allerdings mit einer gewissen Unsicherheit belegt.

Beim Vergleich der Ergebnisse standardisierter Kompetenz- und Leistungstests mit den Befunden zu Geschlechtsunterschieden in Schulnoten werden deutliche Unterschiede zu Ungunsten der Jungen sichtbar. Während sich die besseren Lesekompetenzen der Mädchen auch in ihren besseren Deutschnoten widerspiegeln, gelingt es Jungen anscheinend weniger gut, ihren kleinen, in standardisierten Leistungstests abgebildeten Kompetenzvorsprung in Mathematik auch in gute Mathematiknoten umzuwandeln. In der PISA-Stichprobe von 2006 erhielten sie bei gleichen Testleistungen z.B. schlechtere Noten als Mädchen in Mathematik und naturwissenschaftlichen Fächern (d.h. Biologie, Physik; aber bessere Noten im Fach Deutsch; Blossfeld et al., 2009). In den Stichproben der ELEMENT-Studie (Helbig, 2010b) und der LAU-Studie (Lehmann et al., 1997) erhielten Jungen signifikant schlechtere Deutsch- und Mathematiknoten als Mädchen, wenn für ihre Kompetenzen kontrolliert wurde (für ähnliche Befunde siehe z.B. auch IGLU 2001 Valtin, Wagner & Schwippert, 2005; TIMSS 2007 Neumann, Milek, Maaz & Gresch, 2010). Diese Studien weisen darauf hin, dass die Noten der Jungen in Deutschland folglich auf Grundlage von Kompetenz- und Leistungstests in der Regel überschätzt werden, die der Mädchen hingegen unterschätzt. Zu ähnlichen Ergebnissen kommen auch zahlreiche internationale Studien (z.B. Duckworth & Seligman, 2006; Kenney-Benson, Pomerantz, Ryan & Patrick, 2006; Kimball, 1989; Robinson & Lubienski, 2011).

#### **1.4 Fazit**

Der Unterschied zwischen dem Schulerfolg der Mädchen und dem der Jungen variiert in Abhängigkeit davon, ob (1) Bildungsbeteiligung und Zertifikaterwerb, (2) Noten oder (3) Leistungs- und Kompetenztestergebnisse als Indikatoren für Schulerfolg betrachtet werden. Gegenwärtig besuchen deutlich mehr Mädchen als Jungen das Gymnasium und erwerben das

Abitur. Dies geht zu einem großen Teil darauf zurück, dass Noten als entscheidendes Selektionskriterium in Deutschland bei Schulübergängen und Abschlüssen herangezogen werden und Mädchen bessere Durchschnitts- und Deutschnoten erhalten als Jungen. Gleichzeitig fallen die Geschlechtsunterschiede in standardisierten Schulleistungsuntersuchungen – mit Ausnahme der Lesekompetenz bei 15-Jährigen – in Deutschland klein aus. Am Ende der vierten Klasse entsprechen die Unterschiede in der Lesekompetenz zu Gunsten der Mädchen und in der Mathematikkompetenz zu Gunsten der Jungen maximal dem ungefähren Leistungszuwachs von zwei bis vier Monaten. Die 15-jährigen Mädchen hingegen haben einen Rückstand in der mathematischen Kompetenz von etwa vier bis sechs Monaten gegenüber den Jungen aber einen mehr als doppelt so großen Kompetenzvorsprung im Lesen. Trotz des besseren Abschneidens der Jungen in mathematischen Leistungs- und Kompetenztests erreichen sie jedoch häufig keine besseren oder sogar schlechtere Mathematiknoten als die Mädchen. Die Konsistenz und Robustheit dieser Befunde über verschiedene Länder hinweg verdeutlicht die Bedeutung dieses Phänomens (Voyer & Voyer, 2014). Werden in Studien zur Notengebung zusätzlich die Ergebnisse standardisierter Leistungstests berücksichtigt, vergrößern sich tendenziell die Geschlechtsunterschiede zu Gunsten der Mädchen. Diese Diskrepanz zwischen Testergebnissen und Noten kann natürlich auch als Defizit der Mädchen bei der Bearbeitung standardisierter Leistungstests interpretiert werden (z.B. Kenney-Benson et al., 2006). Da Noten jedoch eine wichtigere Rückmeldefunktion für die Schülerinnen und Schüler selbst haben als Testwerte und außerdem in Deutschland aufgrund ihrer Selektionsfunktion zu den Geschlechtsunterschieden in der Bildungsbeteiligung und im Zertifikateerwerb beitragen, werden sie an dieser Stelle als der zentralere Indikator für Schulerfolg angesehen (siehe auch Steinmayr & Spinath, 2008). Die folgenden Kapitel dieser Arbeit widmen sich dem Versuch, die im Vergleich mit Mädchen und mit standardisierten Leistungstestergebnissen schlechteren Noten der Jungen zu erklären.

## 2 Geschlechtsunterschiede in individuellen schulerfolgsrelevanten Merkmalen

In diesem Kapitel wird vorgestellt, welche individuellen Merkmale von Schülerinnen und Schülern Schulnoten in welchem Ausmaß vorhersagen und inwieweit etwaige Geschlechtsunterschiede auf diesen Merkmalen Unterschiede im Schulerfolg zwischen Jungen und Mädchen erklären können. Angelehnt an die zusammenfassende Darstellung von Spinath, Eckert und Steinmayr (2014) zum gleichen Thema sowie dem Überblicksartikel von Hannover und Kessels (2011) werden Mittelwertunterschiede in Intelligenz, Persönlichkeit, Motivation und schulerfolgsrelevantem Verhalten sowie ihre Bedeutung für das bessere Abschneiden der Mädchen berichtet. Aufgrund der noch dürftigen empirischen Studienlage werden Geschlechtsunterschiede in der Variabilität der Merkmale nicht behandelt.

### 2.1 Intelligenz

Auch wenn in der Wissenschaft seit Jahrzehnten lebhaft darüber diskutiert wird, welche Struktur das Konstrukt Intelligenz aufweist, ist die Bedeutung von Intelligenz für Lernen und Schulerfolg unumstritten (z.B. Gottfredson, 1997; Halpern, 2012; Spinath et al., 2014). Einer verbreiteten Definition zufolge kann Intelligenz verstanden werden als „general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience“ (Gottfredson, 1997, S. 13). Dahinter verbirgt sich die Vorstellung, dass es einen (übergeordneten) Generalfaktor der Intelligenz  $g$  gibt, der die Bearbeitung unterschiedlicher intellektueller Aufgaben unterstützt (z.B. Jensen, 1998; Klauer & Leutner, 2010). Die große Bedeutung der allgemeinen Intelligenz für Schulerfolg zeigt sich auch in der Beschreibung des Konstruktes als das „Potenzial eines Menschen [...], Lern- und Bildungsangebote zur Aneignung von Wissen zu nutzen“ (Stern, 2001, S. 163). Empirisch lassen sich wie erwartet positive

Zusammenhänge mit Schulnoten um  $r = .5$  finden (z.B. Neisser et al., 1996; Sternberg, Grigorenko & Bundy, 2001). Eine höhere allgemeine Intelligenz geht mit besseren Schulnoten einher. Die Zusammenhänge zwischen der allgemeinen Intelligenz und Schulerfolg fallen mit bis zu  $r = .8$  allerdings noch deutlich stärker aus, wenn die allgemeine Intelligenz als latenter Faktor modelliert wird und Testergebnisse als Indikatoren des latenten Konstrukts Schulerfolg herangezogen werden (z.B. Calvin, Fernandes, Smith, Visscher & Deary, 2010; Deary, Strand, Smith & Fernandes, 2007; Duckworth et al., 2012; Spengler, Lüdtke, Martin & Brunner, 2013). Aus diesem Grund wird die allgemeine Intelligenz als einer der stärksten individuellen Prädiktoren für Schulerfolg angesehen (z.B. Gottfredson, 1997; Spinath, Spinath, Harlaar & Plomin, 2006). Intelligenz und Schulerfolg stehen dabei in einer Wechselbeziehung. Denn der Besuch eines Gymnasiums stellt einerseits einen Indikator für Schulerfolg dar und fördert andererseits aufgrund der anspruchsvolleren Lernumgebung auch stärker die Intelligenz als der Besuch anderer Schulformen (z.B. Becker, Lüdtke, Trautwein, Köller & Baumert, 2012).

Einen umfassenden Überblick über den Forschungsstand zu Geschlechtsunterschieden in kognitiven Fähigkeiten insgesamt und den damit verbundenen Herausforderungen bietet Halpern (2012). Tests der allgemeinen Intelligenz zeigen keine konsistenten oder bedeutenden Unterschiede zwischen Jungen und Mädchen auf (z.B. Halpern, 2012; Calvin et al., 2010; Jensen, 1998; Deary et al., 2007; Strand, Deary & Smith, 2006; Spinath et al., 2014). Allerdings wurden Intelligenztests auch so konstruiert, dass sie „genderfair“ sind und keine Geschlechtsunterschiede produzieren (z.B. Halpern, 2012; Neisser et al., 1996). In einer repräsentativen englischen Stichprobe konnten Geschlechtsunterschiede in standardisierten Leistungstests entsprechend auch nicht auf Unterschiede in der allgemeinen Intelligenz zurückgeführt bzw. dadurch erklärt werden (z.B. Calvin et al., 2010). Einige Studien mit Noten als Indikator für Schulerfolg deuten stattdessen darauf hin, dass Intelligenz als Suppressor statt als Mediator wirkt und den Zusammenhang zwischen Geschlecht und

Schulerfolg zu Gunsten der Mädchen erhöht (Steinmayr & Spinath, 2008; Duckworth & Seligman, 2006). Diese Studien deuten darauf hin, dass Intelligenztests ähnlich wie Kompetenztests dazu tendieren, die Schulnoten der Jungen zu überschätzen und die der Mädchen zu unterschätzen. Jungen erhalten folglich eher schlechtere Noten, als aufgrund ihres Abschneidens in Leistungs-, Kompetenz- (siehe Kapitel 1.3) und Intelligenztests (siehe oben) zu erwarten gewesen wäre. Bei der Vorhersage von Collegenoten durch den SAT<sup>2</sup> wurde in den USA ein vergleichbares Phänomen festgestellt und „female underprediction effect“ getauft (Hyde & Kling, 2001).

## 2.2 Persönlichkeit

Persönlichkeit ist „die Gesamtheit aller überdauernden individuellen Besonderheiten im Erleben und Verhalten eines Menschen“ (Asendorpf, 2015). Verschiedene Modelle zur Struktur der Persönlichkeit stimmen darin überein, dass sich diese individuellen Besonderheiten auf fünf großen Persönlichkeitsdimensionen abbilden lassen (z.B. McCrae & Costa, 1987; Digman, 1990; Goldberg, 1990). Diese sogenannten Fünf-Faktoren-Modelle stellen somit die verbreitetsten Persönlichkeitsmodelle dar (Spinath et al., 2014; Wiggins & Trapnell, 1997). Die Bezeichnung der Faktoren variiert leicht zwischen verschiedenen Wissenschaftlerinnen und Wissenschaftlern. An dieser Stelle wird auf die ins Deutsche übersetzte Nomenklatur von McCrae und Costa (Ostendorf & Angleitner, 2004) zurückgegriffen, gemäß der die Faktoren Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit genannt werden und sich je wiederum aus jeweils sechs Facetten zusammensetzen.

Neurotizismus, Extraversion und Verträglichkeit weisen eher schwache oder gar keine signifikanten Zusammenhänge mit Schulerfolg auf (z.B. Poropat, 2009; Spinath et al., 2014). Gewissenhaftigkeit und Offenheit hingegen korrelieren leicht positiv mit Schulerfolg (z.B.

---

<sup>2</sup> Der SAT ist ein standardisierter Studieneingangstest, der zu  $r = .8$  mit allgemeiner Intelligenz korreliert (Frey & Detterman, 2004).

Poropat, 2009; Spengler et al., 2013; Trautwein, Lüdtke, Roberts, Schnyder & Niggli, 2009). Dabei erwies sich Offenheit als stärkerer Prädiktor für Leistungstests als für Noten, während Gewissenhaftigkeit nicht zur Vorhersage von Leistungstestergebnissen, sondern nur von Noten beitrug (Spengler et al., 2013). Etwas stärkere Zusammenhänge mit Noten fanden sich außerdem mit Selbstdisziplin, Pflichtbewusstsein und Leistungsstreben, welche drei der sechs Facetten des Faktors Gewissenhaftigkeit darstellen (Chamorro-Premuzic & Furnham, 2003; Duckworth & Seligman, 2005, 2006).

Mädchen und Frauen erzielten höhere Werte als Jungen und Männer auf den Faktoren Neurotizismus, Extraversion, Gewissenhaftigkeit und Verträglichkeit, die Geschlechtsunterschiede fielen insgesamt aber eher klein aus (z.B. Feingold, 1994; Marsh, Nagengast & Morin, 2013; Schmitt, Realo, Voracek & Allik, 2008). Die Größe der Geschlechtsunterschiede auf den Faktoren Gewissenhaftigkeit und Extraversion variierte jedoch zwischen den Facetten, sodass ihre Betrachtung sinnvoller erscheint (Spinath et al., 2014): Frauen und Mädchen erreichten höhere Werte auf den Extraversionsfacetten Herzlichkeit, Geselligkeit, Aktivität und Frohsinn, für Durchsetzungsfähigkeit und Erlebnishunger zeigte sich hingegen das umgekehrte Bild (Costa, Terracciano & McCrae, 2001). Auch auf den Facetten Selbstdisziplin und Pflichterfüllung des Faktors Gewissenhaftigkeit erreichten Frauen und Mädchen höhere Werte als Männer und Jungen. Die Effektgrößen variierten zwischen klein (Costa et al., 2001; Silverman, 2003) und mittel (Duckworth & Seligman, 2005, 2006; Spinath et al., 2014). Ähnliche Unterschiede zu Gunsten der Mädchen liegen auch für das verwandte Konstrukt Selbstregulation vor (z.B. Kuhl & Hannover, 2012; Matthews, Ponitz & Morrison, 2009).

Duckworth und Seligman (2006) konnten zeigen, dass Geschlechtsunterschiede in der Selbstdisziplin den Geschlechtsunterschied in der Durchschnittsnote zu Gunsten der Mädchen zur Hälfte erklärten. Legten Mädchen und Jungen ein ähnliches Maß an Selbstdisziplin an den Tag, erhielten sie auch ähnliche Noten. Auf ähnliche Weise erklärte die höhere

Gewissenhaftigkeit der Mädchen ihre besseren Sprachnoten (De Fruyt, van Leeuwen, Bolle & Clercq, 2008) bzw. ihr besseres Abschneiden in Noten im Vergleich zu standardisierten Leistungstests (Kling, Nofhle & Robins, 2013). Ein ähnlicher Mediatoreffekt wurde in einer deutschen Studie außerdem für den Persönlichkeitsfaktor Verträglichkeit gefunden (Steinmayr & Spinath, 2008).

Insgesamt hängen die hier vorgestellten Persönlichkeitsmerkmale schwach mit Schulerfolg zusammen. Ebenso unterscheiden sich Mädchen und Jungen auf den meisten Merkmalen nur leicht. Die bisherigen Befunde deuten jedoch darauf hin, dass die stärkeren Ausprägungen auf manchen schulerfolgsförderlichen Persönlichkeitsmerkmalen der Mädchen ihre besseren Schulnoten zum Teil erklären können. So scheinen Mädchen hinsichtlich ihrer Noten besonders von ihrer höheren Selbstdisziplin, Gewissenhaftigkeit und Verträglichkeit zu profitieren (siehe auch Hannover & Kessels, 2011; Spinath et al., 2014).

### **2.3 Motivation**

Unter dem Begriff Motivation werden all die psychischen Prozesse und Zustände subsumiert, die dem menschlichen Verhalten Richtung, Intensität und Ausdauer geben (Urhahne, 2008). Aus der entsprechend breiten und vielseitigen Forschungslandschaft werden an dieser Stelle drei motivationale Konstrukte herausgegriffen, die im Schulkontext besonders ausführlich und intensiv untersucht worden sind (Spinath et al., 2014): Erwartungen und Werte (z.B. Eccles et al., 1983; Wigfield & Eccles, 2000; für eine Darstellung des gesamten Erwartungs-Wert-Modell von Eccles et al. siehe Kapitel 3.1) und Zielorientierungen (z.B. Dweck, 1986; Elliot, 1999).

Erwartungen bezeichnen die wahrgenommene Wahrscheinlichkeit, bei einer Aufgabe oder in einem Fach erfolgreich zu sein (expectancies for success; z.B. Wigfield & Eccles, 2000). In empirischen Studien werden Erwartungen häufig über das Fähigkeitsselbstkonzept (FSK), also einer Selbsteinschätzung der aktuellen Fähigkeiten, operationalisiert. Obwohl es

sich bei Erfolgserwartungen und dem FSK strenggenommen um zwei verschiedene Konstrukte handelt und das FSK ursprünglich als Prädiktor der Erfolgserwartung konzipiert worden ist, ist es nicht gelungen, diese empirisch hinreichend voneinander abzugrenzen (z.B. Eccles & Wigfield, 1995; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000).

Der subjektiv empfundene Wert, den eine Aufgabe oder ein Fachs hat (subjective task values), setzt sich aus den drei Komponenten intrinsischer Wert, persönliche Wichtigkeit und Nützlichkeit zusammen (z.B. Eccles, 2007; Eccles & Wigfield, 1995; Wigfield & Eccles, 1992). Diese werden theoretisch noch durch eine Kostenkomponente ergänzt, die alle negativen Aspekte, die mit der Bearbeitung einer Aufgabe verbunden sind, umfasst, jedoch in der psychologischen Forschung seltener empirisch untersucht wurde (Wigfield & Eccles, 2000). Das Attribut „subjektiv“ soll verdeutlichen, dass die Wertschätzung eines Fachs oder einer Aufgabe von den individuellen Erfahrungen, Merkmalen und dem Umfeld der Schülerin oder des Schülers geprägt und damit nicht statisch sind (Eccles, 2007). Die erste Wertkomponente, die persönliche Wichtigkeit (attainment value), bildet ab, wie wichtig es für eine Person ist, bei der zu erledigenden Aufgabe oder in dem entsprechenden Schulfach gute Leistungen zu erbringen (z.B. Eccles, 2007; siehe auch Battle, 1966). Laut Eccles (2007) wird eine Aufgabe oder ein Fach dann als wichtig wahrgenommen, wenn die Beschäftigung damit das Potenzial hat, zentrale positive Aspekte der eigenen Identität auszudrücken oder zu bestätigen. Aufgaben, die zu den persönlichen Wertvorstellungen und Zielen passen und die eigenen Bedürfnisse erfüllen, wird eine hohe persönliche Wichtigkeit beigemessen (Eccles, 2007). Der intrinsische Wert (intrinsic value) einer Aufgabe ergibt sich aus der Freude oder Vorfreude, die bei oder vor der Bearbeitung dieser erlebt wird (z.B. Eccles, 2007). Diese Komponente weist große Nähe mit dem Konstrukt der intrinsischen Motivation bzw. des Interesses oder des Flows (z.B. Deci & Ryan, 1985; Krapp, 2010; Schiefele, 1991) auf (z.B. Eccles, 2007). Die Nützlichkeit für die Zukunft (utility value) resultiert aus der erwarteten Bedeutung einer Aufgabe für das Erreichen zukünftiger Ziele und kommt damit dem Konzept



der extrinsischen Motivation (z.B. Deci & Ryan, 1985) nahe (z.B. Eccles, 2007). Da die drei Wertkomponenten nicht unabhängig voneinander zu betrachten sind, werden sie typischerweise als Faktoren niedrigerer Ordnung konzipiert und in einem übergeordneten Wertkonstrukt unter dem Begriff „Werte“ zusammengefasst (z.B. Eccles & Wigfield, 1995; Steinmayr & Spinath, 2010). Erwartungen und Werte können sowohl domänenspezifisch als auch domänenunspezifisch erfasst werden, wobei die domänenspezifische Anwendung überwiegt (z.B. Spinath et al., 2014; Steinmayr & Spinath, 2010).

Der empirische Zusammenhang zwischen dem FSK und Noten ist mittelstark bis stark und stärker als der zwischen dem FSK und Leistungstestergebnissen (z.B. Möller et al., 2009; Spinath et al., 2014; Steinmayr & Meißner, 2013), der zwischen Werten und Noten gilt als klein bis mittelstark (z.B. Spinath et al., 2014; Steinmayr & Spinath, 2009). Das FSK und Werte können über Intelligenz hinaus einen Beitrag zur Vorhersage von Noten leisten (z.B. Steinmayr & Meißner, 2013; Steinmayr & Spinath, 2009). Jungen schätzen ihre mathematischen Fähigkeiten höher ein als Mädchen und wertschätzen Mathematik stärker als Mädchen, Mädchen hingegen berichten ein höheres verbales FSK und höhere Werte für die verbale Domäne als Jungen (z.B. Durik, Vida & Eccles, 2006; Steinmayr & Spinath, 2008; Wilgenbusch & Merrell, 1999). Während die Unterschiede im mathematischen FSK und in der Wertschätzung von Mathematik in vielen Studien vollständig den Geschlechtsunterschied in der Wahl fortgeschrittener Mathematikurse mediieren (z.B. Eccles et al., 1983; Eccles, Adler & Meece, 1984; Nagy, Trautwein, Baumert, Köller & Garrett, 2006; Watt, 2006), erhielten Jungen in einer deutschen Stichprobe schlechtere Mathematiknoten, wenn das mathematische FSK und die Werte für Mathematik als vermittelnde Variablen ins Modell aufgenommen wurden (Steinmayr & Spinath, 2008). Gemäß der Definition von MacKinnon, Krull und Lockwood (2000) handelt es sich hierbei um einen Suppressoreffekt. Grundsätzlich können Erwartungen und Werte (individuelle) Unterschiede in Mathematiknoten durchaus erklären (z.B. Berndt & Miller, 1990).

Für die sprachlichen Fächer liegen deutlich weniger Studien vor. Jedoch finden sich auch hier Hinweise darauf, dass die günstigeren Ausprägungen der Mädchen im verbalen FSK und ihre höhere Wertschätzung der verbalen Domäne teilweise ihre besseren Noten und Testleistungen erklären (z.B. Niklas & Schneider, 2012; Steinmayr & Spinath, 2008; siehe auch Spinath et al., 2014).

Zielorientierungen beschreiben, welche Ziele Schülerinnen und Schüler in der Schule verfolgen. Häufig wird zwischen Lernzielen (d.h. die eigenen Kompetenzen erweitern wollen), Annäherungsleistungszielen (d.h. die eigenen Kompetenzen demonstrieren wollen) und Vermeidungsleistungszielen (d.h. eigenes Nichtkönnen verbergen wollen) unterschieden (z.B. Elliot, 1999). Ergänzt wird diese Trichotomie durch das Ziel der Arbeitsvermeidung (z.B. Harackiewicz, Durik, Barron, Linnenbrink-Garcia & Tauer, 2008; Nicholls, Patashnick & Nolen, 1985; Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002). Im Gegensatz zu FSK und Werten werden Zielorientierungen häufig domänenunabhängig und bezogen auf Schule im Allgemeinen erfasst (Wirthwein, Sparfeldt, Pinquart, Wegerer & Steinmayr, 2013). Die Zusammenhänge mit Schulnoten fallen schwach positiv für die ersten beiden genannten Zielorientierungen aus und schwach negativ für Vermeidungsleistungsziele und für Arbeitsvermeidung (z.B. Wirthwein et al., 2013). Die Zusammenhänge mit Leistungstestergebnissen sind (mit Ausnahme der Vermeidungsleistungszielorientierungen) noch schwächer (Wirthwein et al., 2013).

In Bezug auf Geschlechtsunterschiede zeigt sich bei den Zielorientierungen noch kein eindeutiges Bild. Vor allem bei den Leistungszielorientierungen ist der Forschungsstand widersprüchlich (Spinath et al., 2014). Bei den Lernzielen gibt es Hinweise darauf, dass Mädchen diese tendenziell etwas stärker verfolgen (z.B. Kenney-Benson et al., 2006; Nie & Liem, 2013; Steinmayr & Spinath, 2008). Jungen scheinen im Gegenzug stärker das Ziel zu verfolgen, Arbeit zu vermeiden (z.B. Freudenthaler, Spinath & Neubauer, 2008; Steinmayr & Spinath, 2008). Diese Tendenz der Jungen erklärte statistisch zu einem Viertel bis einem

Drittel, warum sie bei gleicher Intelligenz schlechtere Deutsch- und Durchschnittsnoten erhielten als Mädchen (Steinmayr & Spinath, 2008; siehe auch Hannover & Kessels, 2011; Spinath et al., 2014).

Insgesamt zeigen die Studien, dass einige motivationale Konstrukte zur Erklärung von Geschlechtsunterschieden in Schulnoten beitragen können. So können die schlechteren Deutsch- und Durchschnittsnoten der Jungen zu einem Teil statistisch dadurch erklärt werden, dass Jungen stärker das Ziel verfolgen, Arbeit für die Schule zu vermeiden, das Fach Deutsch weniger wertschätzen und ein niedrigeres verbales FSK besitzen als Mädchen. Ihr höheres mathematisches FSK und der höhere Wert, den sie Mathematik zuschreiben, scheinen in Deutschland hingegen als Puffer zu funktionieren und Geschlechtsunterschieden zu Ungunsten der Jungen in Mathematik statistisch entgegen zu wirken.

## **2.4 Verhalten**

Der Begriff Verhalten wird an dieser Stelle als Oberbegriff für verschiedene Verhaltensweisen genutzt, die Schulerfolg vorhersagen können und im Englischen unter dem Term „behavioural engagement“ zusammengefasst werden (siehe Fredricks, Blumenfeld & Paris, 2004). Darunter fallen zum einen Verhaltensweisen wie Anstrengung, Ausdauer, Konzentration und gute Mitarbeit im Unterricht (Fredricks et al., 2004), die in einem positiven Zusammenhang mit Schulerfolg stehen (z.B. Archambault, Janosz, Fallu & Pagani, 2009; Lam et al., 2012). Zum anderen gehört dazu auch ein angemessenes Benehmen in der Schule, wie z.B. möglichst wenig den Unterricht stören oder Regeln und Normen befolgen (Fredricks et al., 2004), welches ebenfalls Schulerfolg positiv vorhersagt (z.B. Archambault et al., 2009; DiPrete & Jennings, 2012). Im Folgenden wird zusammengefasst, welche Bedeutung dem Lernverhalten und Benehmen von Schülerinnen und Schülern für die Erklärung von Geschlechtsunterschieden im Schulerfolg zukommt. Die dargestellten Forschungsergebnisse beruhen dabei auf Selbstberichten der Jugendlichen bezüglich ihres

Verhaltens (für die Relevanz der Fremdwahrnehmung von Lernverhalten und Benehmen durch die Lehrkraft siehe Kapitel 3.3.3).

Mädchen schneiden in beiden genannten Verhaltensclustern besser ab als Jungen. Beispielsweise berichteten Jungen in einer Studie aus zwölf Ländern, sich weniger im Unterricht und für die Schule anzustrengen und ein weniger förderliches Lernverhalten zu zeigen als die Mädchen, und diese Unterschiede erklärten teilweise die besseren Schulleistungen der Mädchen (Lam et al., 2012). Ebenso gaben sich Mädchen mehr Mühe mit und investierten mehr Zeit in das Erledigen der Hausaufgaben, was ebenfalls in einem positiven Zusammenhang mit guten Noten steht (z.B. Trautwein, Lüdtke, Kastens & Köller, 2006). Dem gegenüber steht das stärker störende und weniger angepasste Verhalten von Jungen, welches zu den schlechteren Noten der Jungen beiträgt (Kenney-Benson et al., 2006). Auf die Bedeutung von angemessenem Verhalten und Anstrengung für die Schule für das Erreichen guter Noten weisen auch Studien zur Notengebungspraxis (z.B. Brookhart, 1993; McMillan, 2001; Randall & Engelhard, 2010; siehe Kapitel 1.2) oder mit Lehrereinschätzungen des Verhaltens hin (z.B. Downey & Vogt Yuan, 2005; Duckworth et al., 2012; siehe Kapitel 3.3.3).

Verschiedene Autorinnen und Autoren vermuten, dass das beschriebene schulerfolgsrelevante Verhalten auch die in Kapitel 2.2 und Kapitel 2.3 vorgestellten Zusammenhänge zwischen Persönlichkeitsmerkmalen bzw. Motivation und Schulerfolg erklärt (z.B. De Fruyt et al., 2008; Hannover & Kessels, 2011; Spinath et al., 2014). Tatsächlich zeigten Kenney-Benson et al. (2006) auf, dass das stärkere Verfolgen von Lern- statt Leistungszielen erklärte, warum Mädchen ein adaptiveres Lernverhalten zeigten als Jungen, welches dann wiederum mit besseren Noten einherging. Stärkere Lern- statt Leistungsziele und das damit zusammenhängende engagiertere Lernverhalten mediieren den Geschlechtereffekt auf die Mathematiknote vollständig (Kenney-Benson et al., 2006). Außerdem konnte gezeigt werden, dass sich gewissenhafte Schülerinnen und Schüler und

Schülerinnen und Schüler mit einem positiven FSK mehr im jeweiligen Fach anstrebten, als Schülerinnen und Schüler, die weniger gewissenhaft waren oder ihre eigenen Fähigkeiten im entsprechenden Fach geringer einschätzten (Trautwein et al., 2009). Wie erwartet vermittelte die aufgewendete Anstrengung sowohl den positiven Zusammenhang zwischen Gewissenhaftigkeit und Schulnoten als auch den zwischen dem FSK und Schulnoten. Insgesamt deuten die bisherigen auf den Selbstberichten der Schülerinnen und Schüler basierenden Ergebnisse darauf hin, dass die in Kapitel 2.2 und 2.3 beschriebenen Geschlechtsunterschiede in motivationalen Konstrukten und Persönlichkeitsmerkmalen sich in mehr oder weniger lernförderlichem Verhalten äußern, welches wiederum die besseren Noten der Mädchen und schlechteren Noten der Jungen erklärt.

Zusätzlich fanden sich auch positive – wenn auch deutlich kleinere – Effekte von wenig störendem und lernförderlichem Verhalten auf das Abschneiden in standardisierten Leistungstests (Kenney-Benson et al., 2006; DiPrete & Jennings, 2012). Eine andere Studie wies für das verwandte Konstrukt Selbstkontrolle nur einen positiven Effekt auf die Benotung, aber nicht auf Leistungstests nach (Hofer et al., 2012). Ob die stärkeren Effekte auf Noten darauf zurückzuführen sind, dass Lehrkräfte aus pädagogischen Gründen Anstrengung an sich belohnen wollen (z.B. Brookhart, 1993; McMillan, 2001) oder darauf, dass gutes Lernverhalten die Lehrerwahrnehmung der Schülerleistung im Sinne eines Halo-Effekts unbewusst positiv verzerrt (z.B. Jussim & Eccles, 1992; Kaiser, Retelsdorf, Südkamp & Möller, 2013), ist noch nicht abschließend geklärt.

## **2.5 Fazit**

Obwohl die allgemeine Intelligenz als einer der mächtigsten individuellen Prädiktoren für Schulerfolg gilt, können die schlechteren Noten der Jungen in sprachlichen Fächern und die gemessen an ihrem Abschneiden in Leistungstests schlechteren Noten der Jungen in Mathematik nicht durch etwaige Unterschiede in diesem Merkmal erklärt werden. Auch

Persönlichkeitsmerkmale abseits von Gewissenhaftigkeit und Verträglichkeit scheinen keine bedeutende Rolle für die Entstehung von Geschlechtsunterschieden in Schulnoten zu spielen (Spinath et al., 2014). Einige Studien deuten darauf hin, dass Mädchen insgesamt ein besseres Lernverhalten und Benehmen berichten und dies zu einem Großteil erklärt, warum sie in der Schule bessere Noten erreichen als Jungen (siehe Hannover & Kessels, 2011). Dieses äußert sich zum Beispiel in mehr Engagement und Mitarbeit im Unterricht, disziplinierterem und weniger störendem Verhalten und einer höheren schulischen Anstrengungsbereitschaft und geringeren Arbeitsvermeidung der Mädchen. Lern- und Leistungszielorientierungen einzeln betrachtet konnten hingegen bisher wenig zur Erklärung von Geschlechtsunterschieden im Schulerfolg beitragen. Hinsichtlich der Erklärung von Geschlechtsunterschieden in Schulnoten in den einzelnen Schulfächern kommt den fachspezifischen motivationalen Konstrukten FSK und Werte, welche ebenfalls in einem positiven Zusammenhang mit der Anstrengung für das jeweilige Fach stehen, große Bedeutung zu.

Aus dem in diesem Abschnitt und in Kapitel 1.4 zusammengefassten Forschungsstand zu Geschlechtsunterschieden im Schulerfolg und in individuellen Prädiktoren für Schulerfolg lassen sich zahlreiche weitere interessante Fragestellungen für die Forschung ableiten. Die sich anschließenden Kapitel dieser Arbeit widmen sich den folgenden zwei Fragen: Was erklärt die Geschlechtsunterschiede in der fachspezifischen Motivation und im schulischen Lernverhalten und Benehmen und auf welche Weise hängen diese mit den schlechteren Noten der Jungen zusammen? Was trägt neben dem von den Jugendlichen selbstberichteten Lernverhalten und Benehmen noch zu den schlechteren Beurteilungen der Jungen durch Lehrkräfte bei?

### **3 Ausgewählte Ansätze zur Erklärung von Geschlechtsunterschieden in Motivation, Lernverhalten und Noten**

In diesem Kapitel werden drei mögliche Erklärungen der Geschlechtsunterschiede in Motivation, Lernverhalten und Noten entwickelt. Dazu werden zunächst drei theoretische Ansätze zur Erklärung individueller und geschlechtsbezogener Unterschiede in Schulerfolg und in schulischen Wahlen und zur Eindrucksbildung vorgestellt. Auf dieser Grundlage wird herausgearbeitet, wo die Forschungslücken hinsichtlich der Erklärung der schlechteren Noten der Jungen liegen und welchen zusätzlichen Beitrag der jeweilige Ansatz dabei leisten könnte.

Der erste Ansatz, das Erwartungs-Wert-Modell zur Erklärung schulischer Wahlen und Leistungen von Eccles et al. (im Folgenden: Eccles et al.-Modell; z.B. Eccles et al., 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), ordnet die in Kapitel 2.3 vorgestellten, bedeutenden Prädiktoren für Schulerfolg, fachspezifische Erwartungen und Werte, in ein größeres Gefüge von sozialisierenden Umweltfaktoren und individuellen Merkmalen ein. Eccles et al. stellen in ihrem Modell komplexe Annahmen darüber auf, wie Geschlechtsunterschiede in fachspezifischen Erwartungen und Werten entstehen und wie sie Geschlechtsunterschiede in schulischen Leistungen und Wahlen erklären (z.B. Eccles et al., 1983). Viele der angenommenen Zusammenhänge sind noch nicht empirisch geprüft. Es wird dargelegt, welchen neuen Erkenntnisgewinn eine umfassendere Anwendung des Eccles et al.-Modells zur Erklärung der schlechteren Noten der Jungen in der verbalen Domäne leisten könnte.

Der Ansatz der identitätskongruenten Nutzung schulischer Angebote (z.B. Kessels & Hannover, 2004, 2006) führt Geschlechtsunterschiede in Motivation, Lernverhalten und Interessensentwicklung auf Geschlechtsunterschiede in der erlebten Passung zwischen schulischen Angeboten und der eigenen (Geschlechts-)Identität zurück. Gemäß den

Annahmen dieses Forschungsstrangs bevorzugen Jugendliche die schulischen Angebote und Aktivitäten, die sie aufgrund ihres geschlechtsstereotypisierten Images als zur eigenen geschlechtsbezogenen Identität passend wahrnehmen. Geschlecht wird hier nicht wie in den vorigen Kapiteln nur als Personenmerkmal, sondern auch als psychologisches Merkmal mit den Dimensionen Maskulinität und Femininität verstanden (siehe Eckes & Trautner, 2000a). Kapitel 3.2 fasst die theoretischen Annahmen und die bisherigen Befunde zur Erklärung von domänenspezifischen Geschlechtsunterschieden zusammen. Anschließend wird dargelegt, welche noch zu prüfenden Annahmen zur Erklärung domänenunspezifischer Geschlechtsunterschiede hieraus abgeleitet werden können.

Der dritte im Folgenden vorgestellte Erklärungsansatz greift die Bedeutung der von den Lehrkräften wahrgenommenen Anstrengung der Schülerinnen und Schüler für die Notengebung der Lehrkräfte auf. Basierend auf den theoretischen und empirischen Arbeiten zur Wirkung von Stereotypen auf die Eindrucksbildung wird herausgearbeitet, wie verhaltensbezogene Geschlechterstereotype die Wahrnehmung und Erwartungen von Lehrkräften gegenüber dem Lernverhalten und Benehmen von Schülerinnen und Schülern zu Ungunsten der Benotung von Jungen beeinflussen können. Diesem Ansatz liegt das Verständnis von Geschlecht als sozialer Kategorie zugrunde (siehe Eckes & Trautner, 2000a). Er wird in Kapitel 3.3 Stereotypenbasierte Eindrucksbildung bei Lehrkräften näher dargestellt.

### **3.1 Erwartungs-Wert-Modell**

Nicht nur in der Psychologie, sondern auch in anderen Disziplinen wie der Ökonomie wird menschliches Verhalten verbreitet durch Erfolgserwartungen und Werte erklärt. Das in der Pädagogischen Psychologie erfolgreichste und beliebteste Modell dieser Theorienfamilie ist seit einigen Jahrzehnten das Eccles et al.-Modell zur Erklärung schulischer Wahlen und Leistungen (z.B. Eccles et al., 1983; Wigfield & Eccles, 2000). Angesichts der theoretischen



Komplexität des Modells und der Fülle der durch das Modell inspirierten Studien kann in den folgenden Abschnitten nur ein begrenzter Überblick über die Bedeutung des Modells zur Erklärung von Unterschieden in Motivation und Leistung gegeben werden. Der inhaltlichen Fragestellung der Arbeit entsprechend liegt der Fokus hierbei auf der Erklärung von Geschlechtsunterschieden. Abschließend wird aufgezeigt, welcher weitere Beitrag zur Erklärung des differenziellen Schulerfolgs von Mädchen und Jungen auf Grundlage des Eccles et al.-Modells geleistet werden könnte.

### **3.1.1 Theoretische Annahmen**

Das Eccles et al.-Modell wurde in den 1980er Jahren entwickelt, um Geschlechtsunterschiede – aber auch individuelle Unterschiede im Allgemeinen – in schulischen Wahlen, Anstrengung, Ausdauer und Leistungen in der mathematischen Domäne zu erklären (z.B. Eccles et al., 1983; Eccles et al., 1984; Meece, Parsons, Kaczala & Goff, 1982). Es umfasst sowohl Merkmale des Individuums als auch des sozialen Umfelds inklusive der Personen, die ein Kind sozialisieren. Eccles et al. postulieren, dass Unterschiede in Wahlen, Anstrengung, Ausdauer und Leistungen unmittelbar auf Unterschiede in den motivationalen Merkmalen Erwartung und Wertüberzeugungen zurückgeführt werden können (siehe Kapitel 2.3; z.B. Eccles et al., 1983; Wigfield & Eccles, 2000). Eine hohe Erfolgserwartung und eine hohe Wertschätzung erhöhen die Wahrscheinlichkeit, bei einer Aufgabe oder in einem Fach gute Leistungen zu zeigen oder sich in Zukunft wieder für diese oder dieses zu entscheiden. Im Gegensatz zu früheren Erwartungs-Wert-Modellen (z.B. Atkinson, 1957, 1964) geht das Eccles et al.-Modell von einem positiven statt einem inversen Zusammenhang zwischen den beiden Konstrukten aus (z.B. Eccles & Wigfield, 2002). Erwartungen und Werte sollen überdies additiv und nicht als Multiplikationsterm Schulerfolg und Wahlen vorhersagen (siehe Trautwein et al., 2012).

Das Eccles et al.-Modell stellt außerdem Annahmen darüber auf, welche weiteren Merkmale des Kindes und seiner oder ihrer sozialisierenden Umwelt auf welche Weise zur Entstehung von (individuellen oder Geschlechts-) Unterschieden in Erwartungen und Werten beitragen. Dazu gehören das kulturelle Milieu, die Überzeugungen und das Verhalten der Personen, die das Kind sozialisieren, als relativ „stabil“ angenommene Merkmale des Kindes sowie ihre oder seine bisherigen leistungsbezogenen Erfahrungen (z.B. Wigfield & Eccles, 2000). Das kulturelle Milieu umfasst sowohl gesellschaftlich geteilte Annahmen, wie z.B. Geschlechterrollen oder Geschlechterstereotype, als auch demografische Merkmale der Familie des Kindes (z.B. Eccles & Wigfield, 2002). Geschlechterstereotype sind (sozial geteilte) generalisierte Annahmen darüber, welche Merkmale und Eigenschaften Männer und Frauen bzw. Jungen und Mädchen aufweisen (z.B. Fiedler & Bless, 2003, S. 134; Hilton & von Hippel, 1996, S. 240). Geschlechterrollen bezeichnen die damit verbundenen Verhaltenserwartungen (vgl. Bahrdt, 1984, S. 67; für eine ausführlichere Behandlung der Bedeutung von Stereotypen für Schulleistung siehe Kapitel 3.2 und 3.3). Unter den individuellen, als relativ stabil angenommenen Charakteristiken des Kindes werden z.B. die Intelligenz, Persönlichkeit oder das Temperament subsumiert (z.B. Eccles, 2011).

Die beschriebenen Eigenschaften des Kindes und Merkmale der Umwelt beeinflussen die Erwartungen und Werte des Kindes gemäß dem Modell jedoch nicht direkt, sondern in Abhängigkeit davon, wie das Kind sie wahrnimmt, interpretiert und emotional darauf reagiert (z.B. Eccles, 2011; Wigfield & Eccles, 2000). Der Interpretation und emotionalen Reaktion des Kindes kommt, ebenso wie den kurz- und langfristigen Zielen und dem Selbstbild des Kindes, eine vermittelnde oder moderierende Funktion zu (z.B. Eccles et al., 1983). Nur wenn z.B. eine positive Leistungserfahrung auf die eigenen Fähigkeiten und nicht auf Glück o.ä. attribuiert wird, steigen Erfolgserwartungen und Werte für ähnliche Aufgaben oder das entsprechende Fach.

Mit ihrem Erwartungs-Wert-Modell haben Eccles et al. (1983) eine Meta-Theorie geschaffen, die zahlreiche Wissenschaftlerinnen und Wissenschaftlern seit mehr als drei Jahrzehnten zu vielfältigsten Forschungsaktivitäten inspiriert. Verschiedene bedeutende Theorien wie z.B. Attributionstheorien (z.B. Weiner, 1985), die Selbstbestimmungs-Theorie der Motivation (z.B. Deci & Ryan, 1985) oder die Theorie der tatsächlichen und möglichen Selbstkonzepte (z.B. Markus & Nurius, 1986) können in das Eccles et al.-Modell integriert werden (z.B. Platt, 1988; Wigfield & Eccles, 1992, 2000) und vertiefen dadurch den eröffneten Forschungshorizont. Die Komplexität der vielzähligen theoretischen Konstrukte und ihre angenommenen Beziehungen machen eine Überprüfung des gesamten Modells in einer Studie unmöglich.

Das Modell erscheint zunächst als relativ statisch, doch Eccles et al. nehmen an, dass die Ausprägung der verschiedenen Modellelemente, z.B. der subjektiven Werte, in Abhängigkeit der salienten Merkmale der jeweiligen Situation variieren können (z.B. Eccles et al., 1983; Eccles, 2007). Außerdem ist das Eccles et al.-Modell als Kreislauf konzipiert, in dem jede durch das Modell erklärte Leistung oder Wahl als „frühere“ Lernerfahrung erneut die zukünftigen Werte und Erwartungen beeinflusst (z.B. Eccles, 2011; Eccles & Wigfield, 2002). Obwohl entsprechende Annahmen über die Richtung der Zusammenhänge zwischen den Variablen aufgestellt wurden, wurde zu ihrer Überprüfung in der Regel auf Regressionsanalysen oder Strukturgleichungsmodelle zurückgegriffen, welche ohne entsprechendes Studiendesign keine Aussagen über Kausalität erlauben (z.B. Bollen & Pearl, 2013; Hoyle, 1995).

### **3.1.2 Bedeutung der Merkmale des Individuums**

Dem Namen des Modells entsprechend lag der Schwerpunkt der Forschung auf der Bedeutung von Erwartungen und Werten für schulische Leistungen und Wahlen (z.B. Eccles et al., 1984; Nagy et al., 2006; Watt, 2006). Die Kostenkomponente wurde in der

psychologischen Forschung seltener empirisch untersucht (Wigfield & Eccles, 2000). Obwohl sowohl Erwartungen als auch Werte zunächst aufgabenspezifisch konzeptualisiert wurden (z.B. Eccles et al., 1983; Eccles, 2007), wurden sie in der Forschung meistens bezogen auf ein Schulfach erhoben und zu Erklärung der Unterrepräsentanz von Frauen und Mädchen in der mathematischen und naturwissenschaftlichen Domäne genutzt (z.B. Eccles et al., 1983; Watt, 2006; Wigfield & Eccles, 2000). Dabei erwiesen sich Erwartungen als stärkerer Prädiktor für Leistungen als Werte und Werte erwiesen sich wiederum als stärkerer Prädiktor für Wahlen als Erwartungen (z.B. Berndt & Miller, 1990; Eccles et al., 1984; Meece, Wigfield & Eccles, 1990). In Übereinstimmung mit den Kernannahmen des Modells erklärten Geschlechtsunterschiede in Werten und FSK in vielen Studien teilweise die Geschlechtsunterschiede in Kurswahlen und Noten, wobei die meisten Studien sich auf Mathematikkurswahlen konzentrierten (z.B. Eccles et al., 1983; Eccles et al., 1984; Nagy et al., 2006; Steinmayr & Spinath, 2008; Watt, 2006; siehe Kapitel 2.3). Eine deutsche Studie mit einer Gymnasialstichprobe stellte hingegen fest, dass sich die Geschlechtsunterschiede in den Mathematiknote zu Gunsten der Mädchen vergrößerten, wenn das mathematische FSK und Werte als vermittelnde Variablen aufgenommen wurden (Steinmayr & Spinath, 2008).

Die Bedeutung weiterer im Modell enthaltener individueller Prädiktoren ist schwächer empirisch gesichert. Das allgemeine Selbstbild und die allgemeinen kurz- und langfristigen Ziele des Kindes als Prädiktoren für Erwartungen und Leistungen wurden in dieser ursprünglich breiten Form (z.B. Eccles, 2007; Wigfield, 1994; Wigfield & Eccles, 1992) bisher selten empirisch untersucht. Stattdessen wurden hauptsächlich die Zusammenhänge mit dem spezifischeren Konstrukt der Zielorientierungen geprüft und dabei häufig eine Wirkrichtung entgegengesetzt zu den Annahmen des Eccles et al.-Modells festgestellt (z.B. Greene, DeBacker, Ravindran & Krows, 1999; Plante, O'Keefe & Théorêt, 2013; Wigfield, 1994; siehe auch Wigfield & Cambria, 2010; Wigfield & Eccles, 1992). Ebenso ist die Bedeutung von Elementen des allgemeinen Selbstbildes (z.B. Eccles et al., 1983; Greene et

al., 1999) oder der Wahrnehmung und Interpretation der Überzeugungen der Lehrkräfte und Eltern durch die Schülerinnen und Schüler (z.B. Eccles et al., 1983; Dickhäuser & Stiensmeier-Pelster, 2003) für die Erklärung von Geschlechtsunterschieden in Motivation und Leistung noch vergleichsweise wenig untersucht.

Die Bedeutung von Begabung und Schulleistung zu einem früheren Zeitpunkt für zukünftige Leistungen wurde im Rahmen der Forschung zum Eccles et al.-Modell ebenfalls weniger beachtet. Denn Jungen und Mädchen unterscheiden sich zwar in ihrem Schulerfolg, jedoch nicht in ihrer allgemeinen Intelligenz (z.B. Halpern, 2012; siehe Kapitel 2.1). Ihre Relevanz für die Vorhersage von zukünftigem Schulerfolg ist jedoch aufgrund vielzähliger Studienergebnisse unumstritten (für die Bedeutung früherer Noten siehe z.B. Böhnel, 1996; Huang, 2011; Hülür et al., 2011; für die Bedeutung von Intelligenz siehe z.B. Becker et al., 2012; McDermott et al., 2001; Neisser et al., 1996; Sternberg et al., 2001; siehe auch Kapitel 2.1).

### **3.1.3 Bedeutung der sozialisierenden Umwelt**

Dem Eccles et al.-Modell gemäß trägt die sozialisierende Umwelt nur indirekt zu Unterschieden in der Leistung bei. Die bisherigen Forschungsergebnisse zur Rolle von Geschlechterstereotypen und Elterneinschätzungen kindlicher Fähigkeiten stützen einige der im Modell aufgestellten Zusammenhänge. So trugen z.B. stereotype Annahmen der Eltern über eine größere mathematische Begabung von Jungen als von Mädchen dazu bei, dass sie die mathematischen Fähigkeiten ihrer Töchter unterschätzten, was dann wiederum statistisch erklärte, warum die Mädchen ein niedrigeres mathematisches FSK berichteten als die Jungen (Jacobs & Eccles, 1992; Tiedemann, 2000). In einer ähnlichen Studie mit Schülerinnen und Schülern im Sekundarschulalter ließen sich nicht nur Effekte der Stereotype der Eltern auf die Einschätzung der Fähigkeiten und das FSK des Kindes feststellen, sondern diese sagten auch noch Geschlechtsunterschiede in Noten vorher (Jacobs, 1991). Die Bedeutung von

Unterschieden zwischen den elterlichen Einschätzungen der Fähigkeiten der Jungen und denen der Mädchen für die Erklärung von Geschlechtsunterschieden im FSK wurde in vielen weiteren Studien bestätigt (z.B. Fredricks & Eccles, 2002; Frome & Eccles, 1998; Herbert & Stipek, 2005). Auch das Ausmaß, in dem Lehrkräfte dem Stereotyp anhängen, dass Jungen schlechter lesen könnten als Mädchen, ging mit einer negativeren Selbsteinschätzung der Lesefähigkeiten der Jungen einher – für die Mädchen zeigte sich kein Zusammenhang (Retelsdorf, Schwartz & Asbrock, 2015). Die von den Schülerinnen und Schülern verinnerlichteten fachspezifischen Geschlechterstereotype erklärten – vermittelt über das FSK und die Wertschätzung von Mathematik und Sprache – auch, dass Mädchen bessere Sprach- und Jungen bessere Mathematiknoten erhielten (Plante, de la Sablonnière, Aronson & Théorêt, 2013).

Besonders gut dokumentiert ist darüber hinaus die Bedeutung der sozialen Herkunft des Kindes für Schulerfolg (z.B. Baumert, Stanat & Watermann, 2006a; Statistisches Bundesamt, 2013b). Unter der sozialen Herkunft werden familiäre Merkmale wie der Beruf, das Bildungsniveau oder das Einkommen der Eltern verstanden, die auf verschiedene Arten in Indizes zusammengefasst und klassifiziert werden (für einen Überblick siehe Ehmke & Siegle, 2005). Der wohl berühmteste Einzelindikator, der sozioökonomische Status, wird dabei manchmal auch mit der sozialen Herkunft gleichgesetzt (z.B. Bradley & Corwyn, 2002; Sirin, 2005; White, 1982). Verschiedene Metaanalysen, Schulleistungsuntersuchungen und Überblicksartikel wiesen mittelstarke positive Zusammenhänge zwischen Indikatoren der sozialen Herkunft und den verschiedenen Indikatoren für Schulerfolg wie Bildungsbeteiligung, Noten und Kompetenzerwerb nach (z.B. Maaz, Trautwein, Lüdtke & Baumert, 2008; Müller & Ehmke, 2013; OECD, 2013b; Sirin, 2005; White, 1982). Wie vom Eccles et al.-Modell vorhergesagt konnte dieser Zusammenhang durch Unterschiede im FSK und Werten mediiert werden (Steinmayr, Dinger & Spinath, 2012). In Studien zur Erklärung

von Geschlechtsunterschieden im Schulerfolg auf Grundlage des Eccles et al.-Modells wurde die soziale Herkunft bisher jedoch selten berücksichtigt.

### **3.1.4 Zusammenfassung und Forschungsfazit**

Das Eccles et al.-Modell hat vielfältige und intensive Forschungsbemühen stimuliert und sich zur Erklärung von Geschlechtsunterschieden in schulischen Wahlen und Leistungen in zahlreichen empirischen Untersuchungen bewährt. Es wurde jedoch hauptsächlich (und gemäß seiner ursprünglichen Bestimmung, siehe Eccles et al., 1983; Eccles et al., 1984; Meece et al., 1982) zur Erklärung der Unterrepräsentanz von Mädchen in den MINT-Fächern, vor allem im Hauptfach Mathematik, eingesetzt. Außerdem wurden bisher jeweils kleine Teile des sehr komplexen Modells empirisch überprüft. Eine Modellierung aller im Modell postulierten Zusammenhänge erscheint mithilfe gegenwärtiger statistischer Verfahren und Daten (noch) nicht machbar.

Nicht alle der im Modell mit Pfaden dargestellten komplexen möglichen Zusammenhänge und Interaktionen haben von theoretischer oder empirischer Seite gleich viel Aufmerksamkeit erhalten. Der Schwerpunkt der bisherigen Forschung lag auf den motivationalen Kernelementen Erwartungen und Werte, welche zur Erklärung von Geschlechtsunterschieden in Leistungen und Wahlen beitragen. Auch die Bedeutung von fachspezifischen Stereotypen der Eltern und Lehrkräfte für die Entwicklung von Geschlechtsunterschieden im FSK und Leistungen der Schülerinnen und Schüler wurde im Rahmen von Forschung zum Eccles et al.-Modells gestützt. Diese und weitere Zusammenhänge könnten durch Theorien speziell zur Bedeutung von Geschlechterstereotypen für Schülerinnen und Schüler, wie z.B. dem Modell der Interessensentwicklung als Ausdruck der Identitätsregulation (Kessels & Hannover, 2004, 2006), oder zur Bedeutung für die Eindrucksbildung der Lehrkräfte, wie z.B. dem Kontinuum-Modell der Eindrucksbildung (Fiske & Neuberg, 1990), noch stärker vertieft

werden. Ob dem allgemeinen Selbstbild des Kindes sowie dessen kurz- und langfristigen Zielen, Emotionen und Interpretationen ebenfalls die Bedeutung zukommt, die das Eccles et al.-Modell ihnen zuschreibt, ist ebenfalls noch nicht ausreichend geprüft.

Angesichts der Beliebtheit und Bedeutung des Eccles et al.-Modells zur Erklärung der Unterrepräsentanz der Mädchen in den MINT-Fächern erscheint es vielversprechend, das Eccles et al.-Modell auch auf die schlechteren Noten der Jungen, vor allem in der verbalen Domäne, anzuwenden. Dabei könnten neben den in der Forschung zum Eccles et al.-Modell bereits etablierten Konstrukten FSK, Werte und Elterneinschätzungen der kindlichen Fähigkeiten noch weitere Merkmale des Individuums und der sozialisierenden Umwelt mitaufgenommen und untersucht werden. Aufgrund der Geschlechtsunterschiede in der Bildungsbeteiligung erscheint es für deutsche Studien, die nur eine Schulform berücksichtigen, besonders sinnvoll, auch die Intelligenz und soziale Herkunft der Schülerinnen und Schüler miteinzubeziehen. Beides sind bedeutende Prädiktoren für Schulerfolg und könnten angesichts der Geschlechtsunterschiede in der Bildungsbeteiligung und Selektivität für den Gymnasiumsbesuch (siehe Kapitel 1.1) Geschlechtsunterschiede in Schulnoten und ihren Prädiktoren in Deutschland maskieren.

Die Bedeutung von Geschlechterstereotypen für die Erklärung von Geschlechtsunterschieden in Noten und Motivation, die auch in der Forschung zum Eccles et al.-Modell nachgewiesen wurde, soll in zwei separaten Ansätzen genauer beleuchtet werden. Kapitel 3.2 stellt mit dem Modell der Interessensentwicklung als Ausdruck der Identitätsregulation (z.B. Kessels & Hannover, 2004, 2006) eine Theorie vor, die gezielte Vorhersagen darüber ermöglicht, wie und warum Geschlechterstereotype Unterschiede in Motivation, Verhalten und Noten zwischen Jungen und Mädchen auslösen oder verstärken können. In Kapitel 3.3 wird zuerst auf Grundlage des Kontinuum-Modells der Eindrucksbildung (z.B. Fiske & Neuberg, 1990) herausgearbeitet, wie Geschlechterstereotype auf die Eindrucksbildung von Lehrkräften wirken können, und anschließend



zusammengefasst, welche Konsequenzen dies für die Leistungsbeurteilung von Jungen und Mädchen haben kann. Auf diese Weise wird eine theoriebasierte Analyse der Bedeutung von Geschlechterstereotypen für Geschlechtsunterschiede im Schulerfolg möglich, die über die im Eccles et al.-Modell dargestellten Zusammenhänge hinausgeht.

## **3.2 Identitätskongruente Nutzung schulischer Lernangebote**

Im Zentrum des Ansatzes der identitätskongruenten Nutzung schulischer Lernangebote steht das Modell der Interessensentwicklung als Ausdruck der Identitätsregulation (z.B. Kessels & Hannover, 2004, 2006), welches im folgenden Abschnitt vorgestellt und durch die Annahmen der Identity-Based Motivation Theory (z.B. Oyserman & Destin, 2010) verstärkt und ergänzt wird. Anschließend wird zusammengefasst, welchen Beitrag der Ansatz der identitätskongruenten Nutzung schulischer Angebote bisher zur Erklärung fachspezifischer Geschlechtsunterschiede in Motivation, Verhalten, Wahlen und Leistung geleistet hat. Danach wird vorgestellt, welchen möglichen Beitrag er zur Erklärung der schlechteren Noten der Jungen insgesamt (im Vergleich zu ihrem Abschneiden in Leistungs- und Kompetenztests) leisten könnte (Hannover & Kessels, 2011; Kessels, Heyder, Latsch & Hannover, 2014).

### **3.2.1 Theoretische Annahmen**

Kernelement des Modells der Interessensentwicklung als Ausdruck der Identitätsregulation (interest as identity regulation model, im Folgenden IIRM; z.B. Kessels & Hannover, 2004, 2006; Kessels et al., 2014) und der Identity-Based Motivation Theory (im Folgenden IBMT; z.B. Oyserman, 2009b; Oyserman & Destin, 2010) ist die von den Schülerinnen und Schülern wahrgenommene Passung zwischen ihrer (geschlechtsbezogenen) Identität und (schulischen) Aktivitäten und Fächern. Beide Theorien nutzen den Grad der Passung, um individuelle Unterschiede in Motivation, Interesse, Lernverhalten und damit auch Leistung von Schülerinnen und Schülern zu erklären (Kessels et al., 2014). Damit stehen sie in der Tradition der kognitiven Entwicklungstheorien (z.B. Kohlberg, 1966) und der Geschlechterschematheorien (z.B. Bem, 1983; Martin & Halverson, 1981), welche ebenfalls beinhalten, dass Kinder das Wissen über ihre eigene Geschlechtszugehörigkeit und über die Geschlechtskonnotation, die Gegenständen, z.B. einer Puppe, anhaftet, als Leitlinie für ihr eigenes Verhalten nutzen.

Dem IIRM zufolge nutzen Jugendliche ihre schulische Interessenentwicklung, um ihre tatsächliche oder erwünschte Identität dazustellen bzw. zu schaffen (z.B. Hannover & Kessels, 2011; Kessels & Hannover, 2004, 2006; Kessels et al., 2014). Schule und Unterricht bieten dadurch nicht nur die Möglichkeit zum Wissens- und Kompetenzerwerb, sondern auch einen Raum zur Identitätsentwicklung, welche eine wichtige Entwicklungsaufgabe der Adoleszenz darstellt (z.B. Hannover & Kessels, 2011; Kessels & Hannover, 2004; siehe auch Erikson, 1968, 1988; Havighurst, 1948, 1956). Die Identitätsentwicklung in der Adoleszenz beinhaltet u.a. die Entwicklung und Ausdifferenzierung der eigenen geschlechtsbezogenen Identität, also des individuellen Verständnisses dessen, was es bedeutet, ein Mann oder eine Frau zu sein (z.B. Finkenauer, Engels, Meeus & Oosterwegel, 2002). Im weiteren Sinne umfasst die geschlechtsbezogene Identität alle Aspekte der eigenen Identität, die einen Geschlechtsbezug aufweisen, z.B. auch die Selbstzuschreibung maskuliner oder femininer Eigenschaften, Verhaltensweisen und sozialer Rollen (Eckes & Trautner, 2000a).

Gemäß dem IIRM nutzen, engagieren, interessieren und strengen sich Jugendliche vor allem für die schulischen Angebote an, die sie als kongruent mit ihrer geschlechtsbezogenen Identität wahrnehmen und mit der sie diese demonstrieren können (z.B. Kessels & Hannover, 2004, 2006; Kessels et al., 2014). Dafür greifen sie auf sozial geteilte Annahmen über die Bedeutung schulischer Angebote zurück, wie das Image, das einem Fach oder Verhalten anhaftet, oder welche Prototypen ein Fach oder ein Verhalten repräsentieren (z.B. Hannover & Kessels, 2004; Kessels, 2005; Kessels & Hannover, 2004, 2006; Kessels et al., 2014; Taconis & Kessels, 2009). Dadurch, dass Jugendliche sich aktiv den schulischen Angeboten zuwenden, deren Geschlechtskonnotation zu ihrem Geschlecht passt, können sie die soziale Bedeutung, die einem Lernangebot anhaftet, in die eigene Identität integrieren und ihre (sich noch entwickelnde) Identität als Mann oder Frau betonen (z.B. Kessels & Hannover, 2004; Kessels et al., 2014). In diesem Sinne kann die Bevorzugung identitätskongruenter schulischer Angebote, die dazu dient, auf stilisierte Art und Weise die Zugehörigkeit zur

eigenen Geschlechtsgruppe darzustellen, auch als Form des Gender Enactments oder Gender Displays nach Goffman (1976) aufgefasst werden (siehe auch Kapitel 3.3.2). Das Ausmaß der Identitätskongruenz selbst hängt wiederum nicht nur vom biologischen Geschlecht der Jugendlichen ab, sondern auch davon, wie maskulin und feminin die Jugendlichen sich selbst wahrnehmen, was wiederum auch durch die situative Salienz des Merkmals Geschlecht bedingt ist (Kessels & Hannover, 2004).

Auch die IBMT (z.B. Oyserman, 2009b; Elmore & Oyserman, 2012; für eine ausführliche Darstellung siehe z.B. Oyserman & Destin, 2010) postuliert, dass Angebote stärker angenommen werden, wenn sie als identitätskongruent wahrgenommen werden. Im Gegensatz zur IIRM wurde diese Theorie jedoch nicht spezifisch zur Erklärung von Geschlechtsunterschieden im Schulbereich entwickelt und auch auf andere soziale Kategorien wie z.B. Ethnizität und andere Kontexte wie z.B. Gesundheitsförderung (z.B. Oyserman, Fryberg & Yoder, 2007) oder Konsumverhalten (z.B. Oyserman, 2009a) angewendet. Beide Theorien ergänzen sich jedoch sinnvoll: Ebenso wie das IIRM nimmt auch die IBMT an, dass saliente Aspekte der eigenen Identität handlungsleitend sind und als identitätskongruent wahrgenommenes Verhalten bevorzugt wird. Feine Unterschiede zwischen den Theorien finden sich hinsichtlich der Rolle des Individuums. Während das IIRM stärker als die IBMT den aktiven Prozess der Identitätsregulation durch das Individuum selbst betont, geht Oyserman im Rahmen der IBMT stärker von einer automatisch aktivierten Handlungsbereitschaft (action readiness) aus, die hauptsächlich davon abhängt, welcher Aspekt der eigenen Identität gerade salient ist (dynamic construction; Oyserman, 2009b; Oyserman & Destin, 2010; siehe auch Kessels & Hannover, 2004). Zusätzlich bietet die IBMT mit der unterschiedlichen Interpretation von Schwierigkeiten (interpretation of difficulty) noch eine differenziertere Erklärung dafür, warum identitätskongruente Tätigkeiten auch erfolgreicher ausgeführt werden. Bei identitätskongruenten Tätigkeiten werden Schwierigkeiten als Hinweis auf die Bedeutung und Wichtigkeit der Aufgabe verstanden, was

zum Durchhalten motiviert, während sie bei Tätigkeiten, die als identitätsinkongruent wahrgenommen werden, indizieren, dass die Bearbeitung sinnlos sei (z.B. Oyserman & Destin, 2010). Wird die IBMT zur Erklärung von Geschlechtsunterschieden im Schulerfolg angewandt, sagt sie ähnliches vorher wie das IIRM: Schulische Angebote, die als kongruent zur eigenen geschlechtsbezogenen Identität erlebt werden, werden von Jugendlichen bevorzugt und erfolgreicher ausgeführt werden als identitätsinkongruente.

### **3.2.2 Bedeutung einer Geschlechtsstereotypisierung von Schulfächern**

Gemäß den Annahmen des IIRM sollte eine geschlechtsstereotypisierte Wahrnehmung von einzelnen Fächern oder schulischen Aktivitäten und der daraus resultierende Grad der Passung zur eigenen geschlechtsbezogenen Identität Geschlechtsunterschiede in Motivation, Lernverhalten und Interessensentwicklung erklären (z.B. Kessels & Hannover, 2004, 2006; Kessels et al., 2014). Verschiedene methodische Verfahren und Zugänge wiesen auf ein geschlechtsstereotypisiertes Image von Schulfächern hin: Mathematik und Naturwissenschaften wurden z.B. stärker als Jungen- statt als Mädchenfächer wahrgenommen, bei Sprachen, Kunst und Musik zeigt sich das umgekehrte Bild (z.B. Hannover & Kessels, 2002; Steffens & Jelenec, 2011). Dies spiegelte sich auch in der Erhebung fachlicher Geschlechterstereotype im Selbstbericht wider: Jugendliche schätzten Frauen auf expliziten Ratingskalen für begabter, kompetenter und interessierter an Sprachen und Kunst ein, während sie Männern im Gegenzug Vorteile im Bereich der Mathematik zuschrieben (z.B. Chatard, Guimond & Selimbegovic, 2007; Steffens & Jelenec, 2011; Steffens, Jelenec & Noack, 2010). Weitere Hinweise auf eine Stereotypisierung von Mathematik und Naturwissenschaften als maskulin ergaben sich aus der wahrgenommenen Diskrepanz zwischen diesen Fächern und Femininität. Jugendliche schätzten Mädchen mit Physik als Lieblingsfach als wenig feminin ein (Kessels, 2005). In einer andern Studie löste die Unterstellung einer Physikbegabung bei Mädchen eine kompensatorische Betonung der

eigenen Femininität aus, um diesen durch das Feedback bedrohten Identitätsaspekt wiederherzustellen (Kessels, Warner, Holle & Hannover, 2008). Bei jüngeren Kindern sind die Hinweise auf eine mit expliziten Maßen erhobene Geschlechtsstereotypisierung von Schulfächern etwas weniger eindeutig (z.B. Ambady, Shih, Kim & Pittinsky, 2001; Passolunghi, Rueda Ferreira & Tomasetto, 2014; Steele, 2003).

Studien mit computerbasierten impliziten Messverfahren, wie z.B. dem Impliziten Assoziationentest (Greenwald, McGhee & Schwartz, 1998), wiesen in alters- und ethnisch heterogenen Stichproben aus verschiedensten westlichen Länder konsistent „MINT-männlich“- und „Sprache-weiblich“-Stereotype nach (z.B. Cvencek, Meltzoff & Greenwald, 2011; Nosek, Banaji & Greenwald, 2002; Nosek & Smyth, 2011; Smeding, 2012; Steffens et al., 2010). Diese werden hier als stärkere (implizite) Assoziation eines Schulfachs mit dem Attribut „männlich“ als mit dem Attribut „weiblich“ (oder umgekehrt) definiert und durch Unterschiede in Reaktionszeiten oder Fehlerraten bei verschiedenen Reaktionsaufgaben operationalisiert. Das Adjektiv „implizit“ kann sich sowohl auf die Assoziation, ihre Entstehung oder ihre Messung beziehen, wobei letzteres Verständnis überwiegt (Fiedler, Messner & Bluemke, 2006; Steffens et al., 2010) und auch dieser Arbeit zugrunde liegen soll. Darüber, ob es sich bei implizit und explizit erfassten Stereotypen um unterschiedliche Konstrukte handelt (Greenwald et al., 1998; Wilson, Lindsey & Schooler, 2000), oder ob sie ein Konstrukt darstellen und etwaige Unterschiede zwischen impliziten und expliziten Messungen auf eine unterschiedlich große Manipulierbarkeit der Antworten zurückgehen (Akrami & Ekehammar, 2005), besteht ebenfalls noch keine Einigkeit (z.B. Steffens et al., 2010). Trotz der damit verbundenen Diskussion über die und Forschung zur Reliabilität, Konstruktvalidität und Manipulierbarkeit impliziter Messverfahren im Allgemeinen (z.B. Blanton, Jaccard, Gonzales & Christie, 2006; Cvencek, Greenwald, Brown, Gray & Snowden, 2010; Fiedler et al., 2006; LeBel & Paunonen, 2011; Steffens, 2004) haben sich implizite

Maße zur Vorhersage von Verhalten bewährt (für eine Übersicht siehe z.B. Greenwald, Poehlman, Uhlmann & Banaji, 2009).

Gemäß dem IIRM sollte die Geschlechtsstereotypisierung von Schulfächern Geschlechtsunterschiede in Motivation, Lernverhalten und Interessensentwicklung in dem jeweiligen Fach erklären (z.B. Kessels et al., 2014). Die Ergebnisse verschiedener Studien – besonders auf Grundlage impliziter Maße – untermauerten theoriekonform die Bedeutung der Passung zwischen der eigenen Identität und geschlechtsstereotypisierten Schulfächern für die Erklärung von fachspezifischen Geschlechterdisparitäten. So deckten z.B. Steffens et al. (2010) auf, dass je stärker Frauen Mathematik implizit mit männlich assoziierten (und damit als umso weniger passend zu ihrem eigenen Geschlecht), umso schlechter war ihr mathematisches FSK, ihre Leistung und umso geringer war ihre Intention, einen fortgeschrittenen Mathematikkurs zu wählen (für ähnliche Ergebnisse siehe auch Kiefer & Sekaquaptewa, 2007; Lane, Goh & Driver-Linn, 2012; Nosek et al., 2002; Nosek & Smyth, 2011; Steffens & Jelenec, 2011). Eine zum eigenen Geschlecht inkongruent erlebte Stereotypisierung von Schulfächern und Themen sagte außerdem negativere Einstellungen (z.B. Kessels, Rau & Hannover, 2006; Nosek & Smyth, 2011), ein geringeres fachliches Interesse (z.B. Kerger, Martin & Brunner, 2011) und weniger Engagement in dem jeweiligen Fach vorher (z.B. Nosek & Smyth, 2011; Steffens et al., 2010). Dabei erwiesen sich insgesamt die implizit erfassten Assoziationen als vorhersagemächtiger als die Angaben auf expliziten Maßen (z.B. Kiefer & Sekaquaptewa, 2007; Lane et al., 2012; Nosek & Smyth, 2011; Steffens et al., 2010).

### **3.2.3 Bedeutung einer Geschlechtsstereotypisierung von Schule und Lernen**

Der fächerübergreifend geringere Schulerfolg der Jungen wirft die Frage auf, ob dieser durch eine Stereotypisierung von Schule und schulischem Lernen insgesamt als feminin erklärt werden könnte (Hannover & Kessels, 2011; Kessels et al., 2014). Denn die aus einer

Stereotypisierung von Schule und Lernen als feminin resultierende geringere Passung zur Identität maskuliner Jungen könnte gemäß dem IIRM erklären, warum Jungen weniger lernförderliches Verhalten zeigen als Mädchen und insgesamt schlechtere Noten bekommen (Hannover & Kessels, 2011; Kessels et al., 2014). Ob dem so ist, wurde bisher noch nicht gezielt empirisch untersucht. Deshalb wird im Folgenden zusammengefasst, inwiefern indirekte Hinweise die These der Stereotypisierung von Schule und schulischem Lernen als feminin stützen.

Wie in den Medien diskutiert könnte der hohe Frauenanteil unter den Lehrkräften von mehr als 70% an den allgemeinbildenden Schulen in Deutschland (für das Schuljahr 2012/2013: z.B. Statistisches Bundesamt, 2014) zu einer Wahrnehmung von Schule als feminin führen (siehe auch Brophy, 1985; Huyge, van Maele & van Houtte, 2014; Martino, 2008). In diesem Fall sollten Grundschulen aufgrund ihres höchsten Frauenanteils (2012/2013: 87%) als femininer wahrgenommen werden als Gymnasien (2012/2013: 58%). Dass die Unterschiede im Schulerfolg zu Gunsten der Mädchen in der Sekundarstufe tendenziell größer ausfallen als im Grundschulbereich (siehe Kapitel 1.4), spricht jedoch direkt gegen die Vermutung, dass das Geschlechterverhältnis unter den Lehrkräften zu einer geschlechtsstereotypisierten Wahrnehmung von Schule beitragen könnte. Untermauert wird dieses Argument durch die vielen Studien, die auf Individual- und Schulebene keine systematischen Effekte des Geschlechts der Lehrkraft auf die Motivation, die Noten, die erworbenen Kompetenzen und die Schulempfehlungen der Jungen gefunden haben (z.B. Brophy, 1985; Carrington, Tymms & Merrell, 2008; Cho, 2012; de Zeeuw et al., 2014; Driessen, 2007; Ehrenberg, Goldhaber & Brewer, 1995; Helbig, 2010a, 2010b, 2012; Lam, Tse, Lam & Loh, 2010; Marsh, Martin & Cheng, 2008; Neugebauer, 2011; Neugebauer, Helbig & Landmann, 2011).

Hinweise auf eine Stereotypisierung von schulischem Engagement und Lernverhalten als feminin liefert hingegen die Forschung zu Geschlechterstereotypen, Geschlechterrollen



und dem geschlechtsbezogenen Selbstkonzept von Kindern und Jugendlichen. Wurden z.B. die Stereotype und Metastereotype von Grundschulkindern erfragt, gaben diese an, dass Mädchen sich in der Schule mehr anstrengen, besser benehmen und erfolgreicher sind als Jungen, und dass das auch Erwachsene denken (Hartley & Sutton, 2013). Ein ähnlich negatives Bild über Jungen in der Schule zeigte sich auch in den Antworten deutscher Jugendlicher, welche die Gruppe der Jungen in der Schule häufiger mit negativen Merkmalen wie faul und störend beschrieben als die Gruppe der Mädchen (Latsch & Hannover, 2014). Auch Lehrkräfte schätzten die Gruppe der Mädchen als gewissenhafter und verträglicher ein (Baudson & Preckel, 2013) und sprachen von den „compliant girls and troublesome boys“ (Jones & Myhill, 2004). Die stereotype Annahme, dass Mädchen fleißig seien und sich für die Schule anstrengten, beeinflusste auch die Attributionen von Jugendlichen, Eltern und Lehrkräften. Alle drei Gruppen führten Schulerfolg von Mädchen stärker auf Anstrengung zurück und etwaige Erfolge von Jungen stärker auf Begabung (z.B. Bornholt & Möller, 2003; Lightbody, Siann, Stocks & Walsh, 1996; McClure et al., 2011; Rätty, Vänskä, Kasanen & Kärkkäinen, 2002).

Eine Stereotypisierung von Schule und Lernen als feminin kann außerdem daraus resultieren, dass die weibliche Geschlechtsrolle im allgemeinen als besser zur Schülerrolle passend wahrgenommen wird als die Jungenrolle (z.B. Beaman, Wheldall & Kemp, 2006; Brophy, 1985; Gold & Reis, 1982; Kedar-Voivodas, 1983; Mickelson, 1989; Orr, 2011). Dieser Argumentation zufolge würden Mädchen so sozialisiert, dass sie sich anpassen, unterordnen, vernünftig und fleißig sind. Solches, mit der weiblichen Geschlechtsrolle kongruentes Verhalten sei auch idealtypisches Schülerverhalten und würde mit guten Noten belohnt (Takei, Johnson & Clark, 1998; Beaman et al., 2006; Younger, Warrington & Williams, 1999). Eine Längsschnittstudie, die unabhängig vom Geschlecht des Kindes positive Effekte einer Sozialisation im Sinne der weiblichen Geschlechtsrolle im

Kindergartenalter auf die späteren Schulnoten feststellte, stützt diese Argumentation (Orr, 2011).

Auf eine mögliche Stereotypisierung von schulischem Lernen als feminin weisen auch einige Fragebogen-Items zur Erfassung des geschlechtsbezogenen Selbstkonzepts von Jugendlichen hin. Das geschlechtsbezogene Selbstkonzept wird verstanden als die Verknüpfung der eigenen Identität mit Attributen, die als typisch für ein Geschlecht und untypisch für das andere Geschlecht gelten (z.B. Eckes & Trautner, 2000b)<sup>3</sup>, und häufig in den vier Skalen sozial erwünschte und sozial unerwünschte Femininität bzw. Maskulinität zusammengefasst werden (z.B. Krahé, Berger & Möller, 2007; Spence, Helmreich & Holahan, 1979). Dabei stellt „fleißig“ ein Item zur Erfassung sozial erwünschter Femininität im Jugendalter dar, während „faul“ und „unordentlich“ Items zur Erfassung sozial unerwünschter Maskulinität im Jugendalter sind (Krahé et al., 2007). Gleichzeitig berichteten qualitative, vor allem britische Studien, dass unter den sogenannten „lads“<sup>4</sup> schulerfolgshinderliches Verhalten, wie z.B. sich nicht anstrengen oder den Unterricht stören, als besonders männlich und maskulin gilt, während schulerfolgsförderliches Verhalten als typisches Mädchenverhalten bezeichnet und abgewertet wird (z.B. Epstein, 1998; Francis, 1999; Jackson, 2002, 2003; Jackson & Dempster, 2009). Auch stand eine hohe Ausprägung der sozial unerwünschten Maskulinität – nicht jedoch der sozial erwünschten Maskulinität – in einem ungünstigen Verhältnis zu Schulerfolg, da sie mit einer negativeren Einstellung zum schulischen Hilfesuchen einherging (Kessels & Steinmayr, 2013).

Zu guter Letzt können die stereotypenkonformen Unterschiede in Selbstberichten zum Lernverhalten, zum Engagement, zur Anstrengung und zum Benehmen im Unterricht (siehe Kapitel 2.4) die Stereotypisierung von schulischer Anstrengung und guter Mitarbeit als weiblich und feminin aufrechterhalten oder verstärken. Hierbei handelt es sich jedoch um ein

---

<sup>3</sup> Diesem weiten Verständnis folgend werden die Begriffe geschlechtsbezogenes Selbstkonzept und geschlechtsbezogene Identität in dieser Arbeit synonym verwendet.

<sup>4</sup> Laddishness bezeichnet eine Spielart von Maskulinität unter Jugendlichen. Ein Lad lässt sich im Deutschen am ehesten mit der Bezeichnung Macho charakterisieren (siehe auch Francis, 1999).

„Henne-Ei“-Problem, da sowohl die Geschlechtsunterschiede im selbstberichteten Lernverhalten eine geschlechtsstereotypisierte Wahrnehmung von Lernen befördern können, gleichzeitig aber auch das feminine Image dieses Verhaltens dazu führen kann, dass sich die Geschlechtsunterschiede im Selbstbericht zeigen (für eine Zusammenfassung der Herausforderungen, die mit der Bestimmung der Akkuratheit von Stereotypen verbunden sind, siehe z.B. Schneider, 2004).

Es lässt sich folglich festhalten, dass viele Studien indirekte Hinweise darauf liefern, dass Schule und Lernen insgesamt als feminin stereotypisiert wahrgenommen werden. Dies könnte gemäß dem IIRM eine mögliche Erklärung für den geringeren Schulerfolg der Jungen sein (Hannover & Kessels, 2011; Kessels et al., 2014).

### **3.2.4 Zusammenfassung und Forschungsfazit**

Viele Studien zeigen, dass die Wahrnehmung von MINT-Fächern als maskulin und von Sprachen und Kunst als feminin Geschlechtsunterschiede in den Wahlen, der Motivation, dem Engagement und der Leistung in dem jeweiligen Fach erklärt (z.B. Nosek & Smyth, 2011; Steffens et al., 2010). Der domänenübergreifend geringere Schulerfolg der Jungen (siehe Kapitel 1.4) und die große Bedeutung domänenunspezifischer Konstrukte für Schulerfolg, wie z.B. Lernverhalten und Mitarbeit (siehe Kapitel 2.4), legen nahe, den Ansatz der identitätskongruenten Nutzung schulischer Angebote auch auf Schule und schulisches Lernen insgesamt anzuwenden (Hannover & Kessels, 2011; Kessels et al., 2014). Die im vorigen Kapitel dargestellten Forschungsarbeiten zu Geschlechterstereotypen und –rollen und Männlichkeitsvorstellungen sowie die deutlichen Geschlechtsunterschiede im selbstberichteten schulischen Lernverhalten deuten an, dass auch Schule und Lernen insgesamt eine feminine Konnotation anhaftet. Eine gezielte Überprüfung einer Stereotypisierung von Schule und schulischem Lernen als feminin steht jedoch noch aus. In diesem Fall sollten Schule und schulisches Lernen als weniger gut kompatibel zu der Identität

von Jungen wahrgenommen werden, vor allem von den Jungen, die sich selbst als sehr maskulin beschreiben. Entsprechend sollten (maskuline) Jungen sich theoriegemäß stärker von der Schule abwenden, weniger lernförderliches Verhalten zeigen und in der Folge schlechtere Noten bekommen, vor allem im ebenfalls feminin stereotypisierten Fach Deutsch. Diese vermuteten, differenziellen Zusammenhänge zwischen dem Geschlecht, dem geschlechtsbezogenen Selbstkonzept der Jugendlichen, einer möglichen geschlechtsstereotypisierten Wahrnehmung von Schule und Schulerfolg gilt es noch empirisch zu testen. Dabei erscheint es sinnvoll, die Geschlechtsstereotypisierung mit impliziten Maßen zu erheben, da sich diese bei der Erklärung von fachspezifischen Geschlechtsunterschieden bereits als vorhersagemächtiger erwiesen haben als explizite Maße. Außerdem sollte zwischen sozial erwünschter und unerwünschter Maskulinität differenziert werden, da besonders Letztere mit schulischen Anforderungen im Konflikt zu stehen scheint.

### **3.3 Stereotypenbasierte Eindrucksbildung bei Lehrkräften**

Das Stereotyp des faulen Jungens und des fleißigen Mädchens kann – wie in den vorigen Abschnitten dargelegt – Schule und schulischem Engagement und Lernen ein feminines Image geben und damit die Identitäts- und Interessensentwicklung und das schulische Verhalten der Jugendlichen beeinflussen. Gleichzeitig kann es auch auf die Kognitionen der Lehrkräfte wirken und beeinflussen, welche Unterschiede Lehrkräfte zwischen Schülerinnen und Schüler in welchem Ausmaß wahrnehmen. Während in Kapitel 3.2 der Schwerpunkt auf der Bedeutung einer möglichen Geschlechtsstereotypisierung von Schule und schulischem Engagement für Jugendliche selbst lag, steht im Zentrum dieses Kapitels die Wirkung der beschriebenen verhaltensbezogenen Geschlechterstereotype auf die Eindrucksbildung von Lehrkräften. Gemäß dem Verständnis von Geschlecht als sozialer Kategorie (z.B. Eckes & Trautner, 2000a) liegt der Fokus hier folglich auf der Rolle des Beobachtenden bei der Entstehung von Geschlechtsunterschieden. In den folgenden Unterkapiteln wird herausgearbeitet, wie stereotype Annahmen über das schulerfolgsrelevante Verhalten von Mädchen und Jungen bei Lehrkräften ausgelöst werden und ihre Erwartungen bezüglich des Lernverhaltens und Benehmens zu Ungunsten der Jungen verzerren können. Anschließend wird dargelegt, wie allein der Eindruck, dass eine Schülerin oder ein Schüler sich beispielsweise weniger anstrengt oder mehr stört, zu einer schlechteren Benotung und Kompetenzeinschätzung dieser Person durch Lehrkräfte führen kann.

#### **3.3.1 Wirkung von Stereotypen auf die Eindrucksbildung**

Stereotype sind als Schemata im Gedächtnis abgespeichert und können von dort die Informationsverarbeitung in sozialen Interaktionen systematisch beeinflussen (z.B. Brewer, 1988; Fiedler & Bless, 2003; Fiske & Neuberg, 1990; Schneider, 2004). Im Sinne eines Top-Down-Prozesses können sie auf die Enkodierung und Interpretation neuer Informationen, z.B.

über eine fremde Person, einwirken, welche wiederum die Grundlage weiterer Schlussfolgerungen, Entscheidungen und Urteile bilden (Fiedler & Bless, 2003).

Gegenwärtige Stereotype über Jungen und Mädchen in der Schule beinhalten nicht nur fachspezifische Kompetenzunterschiede, sondern auch fachübergreifende Verhaltensunterschiede: So werden Mädchen von Schülerinnen, Schülern und Lehrkräften als fleißig und angepasst charakterisiert und Jungen als faul und störend (z.B. Hartley & Sutton, 2013; Jones & Myhill, 2004; Latsch & Hannover, 2014; siehe auch Kapitel 3.2.3). Das Kontinuum-Modell der Eindrucksbildung (Fiske, Lin & Neuberg, 1999; Fiske & Neuberg, 1990) bietet einen theoretischen Rahmen, um die Wirkung dieser Stereotype auf die Eindrucksbildung von Lehrkräften näher zu untersuchen. Grundannahme des Modells ist, dass sich die Informationsverarbeitung und Eindrucksbildung in sozialen Interaktionen zwischen zwei Polen bewegt. Der erste Pol stellt eine ausschließlich auf sozialen Kategorien basierende und damit stereotypenbasierte Informationsverarbeitung dar, der andere eine Informationsverarbeitung, bei der die Zugehörigkeit zu einer sozialen Kategorie nur ein gewöhnliches individuelles Merkmal einer Person darstellt und wenig Einfluss auf die Eindrucksbildung hat (Fiske & Neuberg, 1990; für ein weiteres Modell, das zwischen diesen beiden Arten der Informationsverarbeitung unterscheidet, siehe Brewer, 1988). Damit ein Stereotyp in der sozialen Interaktion mit einem Individuum wirken kann, muss die Targetperson zunächst als Mitglied der sozialen Kategorie, die mit dem Stereotyp belegt ist, identifiziert werden. Nach dem Kontinuum-Modell der Eindrucksbildung läuft dieser Prozess der Kategorisierung in sozialen Begegnungen augenblicklich dann ab, sobald genügend Hinweisreize vorliegen, um auf die Gruppenzugehörigkeit der Targetperson zu schließen (Fiske & Neuberg, 1990). Mit der Kategorisierung werden kategoriebezogene Kognitionen (Stereotype), Emotionen (Vorurteile) und Verhaltenstendenzen zugänglich (Fiske & Neuberg, 1990). Den Kategorien Geschlecht, Ethnizität und Alter kommt dabei eine Sonderstellung zu, da sie äußerlich sichtbar sind bzw. gemacht werden und innerhalb von Millisekunden

verfügbar sind (Fiske et al., 1999). In experimentellen Studien konnte auch gezeigt werden, dass allein das Verwenden eines männlichen oder weiblichen Vornamens geschlechterstereotype Erwartungen und Beurteilungen auslöst (z.B. Baudson & Preckel, 2013; Condry & Condry, 1976; Dipboye, Fromkin & Wiback, 1975). Wenn die Targetperson weder interessant noch relevant ist, oder es keine Möglichkeit gibt, weitere Informationen über sie zu erhalten (z.B. in experimentellen Settings), bleibt der gewonnene erste Eindruck stark kategorien- also stereotypenbasiert (Brewer, 1988; Fiske & Neuberg, 1990). Wenn die Targetperson hingegen interessant oder persönlich relevant ist – wovon in einem Lehrer-Schüler-Verhältnis auszugehen ist – werden auch ihre individuellen Merkmale beachtet: Je besser sie zur ursprünglichen Kategorisierung passen, desto salienter wird die soziale Kategorie und umso stärker ist ihr Einfluss auf die Eindrucksbildung (z.B. Brewer, 1988; Fiske et al., 1999; Fiske & Neuberg, 1990). Passen die individuellen Merkmale nicht zur ursprünglichen Kategorisierung, kommt es entweder zu einer Re-Kategorisierung oder einer schrittweise erfolgenden Analyse und Integration der individuellen Merkmale der Targetperson („piecemeal integration“). Im letzten Fall beruht der entstandene Eindruck fast ausschließlich auf den individuellen Merkmalen der Targetperson und wenig auf kategoriebezogenen Informationen wie Stereotypen (Fiske & Neuberg, 1990).

Da Stereotype im Erstkontakt automatisch aktiviert werden, jedoch von einem grundsätzlichen Interesse der Lehrkräfte an ihren Schülerinnen und Schülern ausgegangen werden kann, macht besonders die mittlere Phase des Modells eine interessante Vorhersage: Geschlechterstereotype sollten die Bildung erster Eindrücke bei Lehrkräften stärker beeinflussen, wenn die individuellen Merkmale der Schülerin oder des Schülers die automatische Einordnung in die Kategorie „männlich“ oder „weiblich“ bestätigen, als wenn nicht. Dies bedeutet auch, dass die Jugendlichen, deren individuelle Merkmale oder individuelles Verhalten ihre Zugehörigkeit zur Gruppe der Männer oder Frauen betonen,

stärker stereotypenbasierte Erwartungen bei Lehrkräften hervorrufen sollten als die Jugendlichen, deren Männlichkeit oder Weiblichkeit weniger deutlich sichtbar ist.

### **3.3.2 Aktivierung von Geschlechterstereotypen durch Gender Enactment**

Schülerinnen und Schüler können durch ihr Auftreten in sozialen Interaktionen ihre Zugehörigkeit zur Gruppe der Männer oder Frauen hervorheben. Dies kann z.B. über Kleidung, Aussehen oder Verhaltensweisen erfolgen, welche in der Gesellschaft als maskulin oder feminin gelten und verwendet werden, um Männer und Frauen voneinander zu unterscheiden (z.B. Eckes & Trautner, 2000a; Hilliard & Liben, 2010). Solch stilisierte Darstellung des eigenen Geschlechts wird als Gender Enactment oder Gender Display bezeichnet (Goffman, 1976) und unterstützt die Identitätsfindung in sozialen Interaktionen (z.B. Deaux & Major, 1987). Gerade in der Adoleszenz können Jugendliche von Gender Enactment profitieren, da in der Pubertät die Zugänglichkeit des geschlechtsbezogenen Selbstwissens erhöht ist (Finkenauer et al., 2002) und Jugendliche vor der Aufgabe stehen, eine (differenziertere) geschlechtsbezogene Identität zu entwickeln (z.B. Hannover & Kessels, 2011; Kessels & Hannover, 2004; siehe auch Erikson, 1968, 1988; Havighurst, 1948, 1956). Im Schulkontext kann die Bearbeitung dieser Entwicklungsaufgabe, wie in Kapitel 3.2.1 dargestellt, in Unterschieden in der Interessens- und Leistungsentwicklung zwischen Jungen und Mädchen münden (Kessels & Hannover, 2006; Kessels et al., 2014). So kann z.B. auch die Distanzierung der Mädchen von den MINT-Fächern als Gender Enactment angesehen werden, wenn sie dazu dient, die eigene Femininität zu demonstrieren bzw. bei Bedrohung wiederherzustellen (Kessels et al., 2008). Im Gegenzug könnte ein gesteigertes Interesse der Jungen an MINT-Fächern als Inszenierung der eigenen Maskulinität verstanden werden. Aber auch eine bewusste Ablehnung von Schule und schulischem Lernen könnte in manchen Kreisen zur Maskulinitätsinszenierung genutzt werden (z.B. Jackson, 2002, 2003). Darüber hinaus sind jedoch auch viele weitere Möglichkeiten und Arten des Gender Enactments in der



Schule vorstellbar, die keinen direkten Bezug zu Schulerfolg aufweisen und sich z.B. in der Kleidung, dem Gebrauch von Make-Up und Kosmetika, der Körperhaltung oder Ähnlichem manifestieren (z.B. Eckes & Trautner, 2000a; Hilliard & Liben, 2010).

In dem gesamten Prozess kommt den Peers eine zentrale Rolle zu. Sie sind sowohl das Publikum als auch die Kritiker und Kritikerinnen der „Aufführung“ und bestimmen die Bewertungsmaßstäbe (vgl. Kessels et al., 2008). Ihre Bedeutung spiegelt sich im Zusammenhang mit Beliebtheit und sozialer Akzeptanz wider: So versuchen z.B. Kinder absichtlich, sich typisch männlich oder weiblich zu verhalten, weil sie denken, dass sie das bei ihren Peers beliebt macht (Adler, Kless & Adler, 1992). In der Adoleszenz steigt gemäß der Gender-Intensification-Hypothese der soziale Druck, sich möglichst geschlechtskonform zu verhalten (Hill & Lynch, 1983). Empirisch zeigte sich, dass Jugendliche, die sich als typisch für ihr Geschlecht wahrnehmen, bei ihren Peers beliebter und stärker sozial akzeptiert sind und über einen höheren Selbstwert verfügen als Jugendliche, die sich als weniger typisch für ihr Geschlecht einschätzen (z.B. Egan & Perry, 2001; Jewell & Brown, 2014; Yunger, Carver & Perry, 2004).

Auf Grundlage des Kontinuum-Modells kann die Hypothese aufgestellt werden, dass Gender Enactment von Schülerinnen und Schülern den Einfluss von Geschlechterstereotypen auf die Eindrucksbildung bei Lehrkräften verstärkt. Jungen, die ihre Maskulinität inszenieren, sollten theoriegemäß mit negativeren Erwartungen der Lehrkräfte hinsichtlich ihres Lernverhaltens und Benehmens konfrontiert sein als Mädchen, da die Gruppe der Jungen in der Schule zurzeit – vereinfacht gesagt – als faul und weniger angepasst wahrgenommen wird (z.B. Hartley & Sutton, 2013, siehe Kapitel 3.2.3). Dies sollte theoriegemäß auch dann der Fall sein, wenn die Verhaltensweisen, die zum Gender Enactment genutzt werden, selbst in keinem direkten Zusammenhang mit Schulerfolg stehen.

### 3.3.3 Bedeutung des zugeschriebenen Verhaltens für Schulerfolg

Während in Kapitel 2.4 die Bedeutung des von den Schülerinnen und Schülern selbstberichteten Lernverhaltens und Benehmens für Schulerfolg zusammengefasst wurde, stellt dieser Abschnitt die Bedeutung des den Schülerinnen und Schülern von den Lehrkräften *zugeschriebenen* Lernverhaltens und Benehmens für die Notengebung und Kompetenzbeurteilung vor. Dazu wird auf die Ergebnisse von Vignettenexperimenten und Szenariostudien zur Notengebungspraxis, Feldstudien zur Bedeutung des von den Lehrkräften eingeschätzten Lernverhaltens und Benehmens und auf Studien mit dem „Simulierten Klassenraum“ zurückgegriffen.

Vignettenexperimente und Szenariostudien zeigen deutlich und konsistent die potentielle Bedeutung von Annahmen über das Lernverhalten<sup>5</sup> für die Notengebung auf: Lasen Lehrkräfte, dass fiktive Schülerinnen und Schüler sich anstrebten und die Klassenregeln befolgten, bewerteten sie diese bei gleichen Leistungen besser, als fiktive Schülerinnen und Schüler, über die die Lehrkräfte lasen, dass sie sich weniger anstrebten und im Unterricht durch unangemessenes oder störendes Verhalten negativ auffielen (z.B. Brookhart, 1993; Randall & Engelhard, 2010; Sun & Cheng, 2014). Besonders schwache Schülerinnen und Schüler wurden deutlich positiver benotet, wenn sie sich anstrebten und angemessen verhielten (Randall & Engelhard, 2010). Selbstberichte von Lehrkräften bestätigen diese Praxis der Notengebung (z.B. Cizek et al., 1995; McMillan, 2001; McMillan et al., 2002; siehe auch Kapitel 1.2). Auch in Feldstudien war das Lernverhalten und Benehmen, das Lehrkräfte den Schülerinnen und Schülern attestierten, ein wichtiger Prädiktor ihrer Noten (z.B. Bennett, Gottesman, Rock & Cerullo, 1993; DiPrete & Jennings, 2012; Downey & Vogt Yuan, 2005; Jussim & Eccles, 1992; McDermott et al., 2001; siehe auch Jussim, Eccles & Madon, 1996). Dabei leistete die Lehrereinschätzung der Anstrengung einen

---

<sup>5</sup> Lehrkräfte können Annahmen sowohl bezüglich des Lernverhaltens als auch bezüglich der Leistungsfähigkeit von Schülerinnen und Schülern haben. Letztere haben unter dem Schlagwort „Lehrererwartungseffekte“ große Aufmerksamkeit erfahren. Auf sie wird an dieser Stelle nicht näher eingegangen. Für einen Überblick über ihre Bedeutung für Schulerfolg siehe z.B. Jussim und Harber (2005) oder Madon, Willard, Gyll und Scherr (2011).

eigenen Beitrag zur Vorhersage der Noten, der über die von den Schülerinnen und Schülern selbst berichtete Anstrengung hinaus ging (Jussim et al., 1996). Das durch die Lehrkräfte eingeschätzte Lernverhalten und Benehmen erklärte – ähnlich wie das von den Schülerinnen und Schülern selbstberichtete Lernverhalten (siehe Kapitel 2.4) – auch, warum Jungen bei gleichen früheren Leistungen oder Kompetenzen schlechter beurteilt wurden als Mädchen (DiPrete & Jennings, 2012; Downey & Vogt Yuan, 2005).

Das von den Lehrkräften wahrgenommene oder zugeschriebene Lernverhalten wirkt sich jedoch nicht nur auf die Notengebung aus, sondern kann auch die Kompetenz- und Leistungsbeurteilung beeinflussen. Kaiser et al. (2013) wiesen in einer Querschnittsstudie und zwei „Simulierter Klassenraum“-Experimenten nach, dass Engagement von Schülerinnen und Schülern zu einer Überschätzung ihrer Leistungen und Kompetenzen durch die Lehrkräfte führt, auch unter Kontrolle der tatsächlichen Leistung. Gleichzeitig konnten sie zeigen, dass eine hohe Leistung zu einer Überschätzung des Engagements der Schülerin oder des Schülers führt – auch wenn durch experimentelle Kontrolle das tatsächliche Engagement und die tatsächliche Leistung unkorreliert waren. Sie interpretierten diese beidseitigen Wirkzusammenhänge als Halo-Effekt, bei dem jeweils ein positives Merkmal auf die Wahrnehmung des anderen Merkmals positiv wirkt (Kaiser et al., 2013). Es ist folglich zu erwarten, dass sowohl eine Überschätzung des tatsächlichen Engagements als auch eine Überschätzung der tatsächlichen Leistung sich in besseren Noten niederschlagen (siehe auch Kapitel 1.2).

Insgesamt zeigen die Studien auf, dass nicht nur das selbstberichtete Verhalten von Schülerinnen und Schülern ihren Schulerfolg vorhersagt, sondern auch das Verhalten, dass Lehrkräfte ihren Schülerinnen und Schülern zuschreiben. Es reicht aus, dass Lehrkräfte lediglich *lesen*, dass eine Schülerin oder ein Schüler sich angestrengt oder gestört hat, oder dass bei ihnen der *Eindruck* entstanden ist, dass eine Schülerin oder ein Schüler sich wahrscheinlich anstrengt oder eben nicht, um unterschiedliche Benotungen und

Kompetenzeinschätzungen auszulösen. Dabei muss der entstandene Eindruck der Lehrkräfte nicht mit der Selbstwahrnehmung der Schülerinnen und Schüler übereinstimmen, was sich in den niedrigen Zusammenhängen zwischen beiden Maßen zeigt (Kaiser et al., 2013, Studie 1; siehe auch Jussim & Eccles, 1992).

### **3.3.4 Zusammenfassung und Forschungsfazit**

Lehrkräfte benoten Schülerinnen und Schüler, die sich angestrengt haben oder von denen sie sie den Eindruck haben, dass sie sich angestrengt haben, besser als Schülerinnen und Schüler, von deren Lernverhalten sie einen negativeren Eindruck haben. Ähnliches gilt für angepasstes, wenig störendes Verhalten in der Schule. Angesichts dieser Relevanz des Lernverhaltens und Benehmens für die Notengebung stellt sich die Frage, welche Prozesse dazu beitragen, dass bei Lehrkräften der Eindruck entsteht, dass eine Schülerin oder ein Schüler sich nicht anstrengt. Einen potentiell mächtigen Einfluss auf die Eindrucksbildung stellen Geschlechterstereotype dar. Da aktuelle Geschlechterstereotype Jungen ein schlechteres Lernverhalten und Benehmen in der Schule insgesamt zuschreiben als Mädchen, sollte eine Aktivierung dieser Stereotype bei Lehrkräften negativere Erwartungen bezüglich des Lernverhaltens und Benehmens von Jungen im Vergleich zu dem von Mädchen auslösen. Aus dem Kontinuum-Modell der Eindrucksbildung (Fiske & Neuberg, 1990) lässt sich ableiten, dass Gender Enactment den Einfluss dieser verhaltensbezogener Geschlechterstereotype auf die Eindrucksbildung von Lehrkräften zu Ungunsten der Jungen noch zusätzlich verstärken kann. Für die Forschung schließt sich daran der Auftrag an, empirisch zu testen, ob Gender Enactment von Jugendlichen bei Lehrkräften tatsächlich verhaltensbezogene Geschlechterstereotype aktivieren kann, welche sich dann wiederum in einem negativeren Eindruck der Lehrkräfte bezüglich des Lernverhaltens und Benehmens der Jungen niederschlagen.

## 4 Fragestellungen, Ziele und Überblick der Studien

Im Kapitel 3 wurden drei theoretische Ansätze zur Erklärung von Geschlechtsunterschieden in Motivation, Lernverhalten und Schulerfolg vorgestellt und zusammengefasst, inwieweit Ergebnisse empirischer Studien die theoretischen Annahmen stützen. Auf dieser Grundlage wurde herausgearbeitet, wo jeweilige Forschungslücken hinsichtlich der Erklärung der schlechteren Noten der Jungen liegen und welchen weiteren Beitrag der jeweilige Ansatz zur Schließung dieser Lücken leisten könnte. Die daraus abgeleiteten Fragestellungen lassen sich folgendermaßen zusammenfassen:

- (1) Können bedeutende Variablen des Eccles et al.-Modells, d.h. FSK, Werte, Elterneinschätzung der kindlichen Fähigkeiten, Intelligenz, Bildungsstand der Eltern und die früheren Leistungen, die Geschlechtsunterschiede nicht nur in den Mathematik- sondern auch in den Deutschnoten erklären?
- (2) Ist Schule implizit feminin stereotypisiert und welche Rolle spielt dies für den differenziellen Schulerfolg von Jungen und Mädchen?
- (3) Kann Gender Enactment von Jugendlichen bei Lehrkräften stereotypisierte Erwartungen zu Ungunsten des Lernverhaltens und des Benehmens von Jungen auslösen?

Um diese Fragestellungen empirisch zu untersuchen, wurden drei Studien entwickelt und durchgeführt. Die Studien 1 und 2 ziehen Schulnoten als Indikator für Schulerfolg heran, da Schulnoten eine wichtigere Rückmeldefunktion für die Schülerinnen und Schüler selbst haben als Leistungstestergebnisse und außerdem in Deutschland aufgrund ihrer Selektionsfunktion zu den Geschlechtsunterschieden in der Bildungsbeteiligung und im Zertikateerwerb beitragen. Besonderes Augenmerk lag in beiden Studien auf den Deutschnoten, da sich hier besonders deutlich und konsistent Unterschiede zu Ungunsten der Jungen zeigen. Studie 3 untersuchte die Entstehung des Eindrucks bei Lehrkräften, dass eine Schülerin oder ein

Schüler ein schlechteres Lernverhalten zeigt und sich unangepasster benimmt, da Lehrkräfte einen solchen Eindruck bei ihrer Notengebung negativ berücksichtigen. Die gemeinsame Grundlage von Studie 2 und 3 bildet die Annahme, dass Schule und schulisches Lernen feminin stereotypisiert wahrgenommen werden und lernförderliches und angemessenes Verhalten in der Schule als typisch für Mädchen, aber nicht für Jungen, gilt. Die folgenden Abschnitte geben einen kurzen Überblick über die Studien.

Die erste Studie (*Beyond expectancies and values: Applying the Eccles et al. model to the underachievement of boys*) untersucht die Frage, ob zentrale Variablen des Eccles et al. Erwartungs-Wert-Modells (z.B. Eccles et al., 1983) Geschlechtsunterschiede in Deutschnoten zu Ungunsten der Jungen genauso gut erklären können wie die Geschlechtsunterschiede in den Mathematiknoten zu Ungunsten der Mädchen. Da sich das Eccles et al.-Modell zur Vorhersage schulischer Leistungen und Wahlen in den maskulin konnotierten MINT-Fächern bereits bewährt hat (z.B. Watt, 2006), wurde in Studie 1 in einer Stichprobe von  $N = 520$  Oberstufenschülerinnen und -schülern geprüft, ob es ähnlich erfolgreich zur Erklärung der vergleichsweise untererforschten schlechteren Noten der Jungen in Sprachen angewendet werden kann. Die Studie geht über die bisherigen Studien hinaus, da neben den motivationalen Merkmalen (FSK, Werte) auch weitere im Modell enthaltene individuelle Merkmale (Intelligenz, frühere Leistung) und Elternvariablen (Bildungsstand, Elterneinschätzung der Fähigkeiten des Kindes) erhoben wurden. Diese wurden als Mediatoren in zwei simultane multiple Mediationsmodelle zur Erklärung der Geschlechtsunterschiede in den Deutsch- und Mathematiknoten aufgenommen und die Ergebnisse zwischen beiden Domänen verglichen.

Nachdem viele Studien zeigen konnten, dass Schulfächer stärker mit dem einen als mit dem anderen Geschlecht assoziiert werden und diese Geschlechtsstereotypisierung motivationale Merkmale, Leistung und Wahlen des entsprechenden Fachs vorhersagt (z.B. Nosek & Smyth, 2011; Steffens & Jelenec, 2011), wurde in Studie 2 (*Is school feminine?*

*Implicit gender stereotyping of school as a predictor of academic achievement*) geprüft, ob ähnliches auch für Schule und schulisches Lernen insgesamt gilt. Mithilfe eines impliziten Messverfahrens, des Go/No-go Association Tasks (Nosek & Banaji, 2001), wurde in einer Stichprobe von  $N = 122$  Schülerinnen und Schüler der neunten Klasse überprüft, ob Schule und schulischem Lernen insgesamt eine feminine Konnotation anhaftet und dies entsprechend der theoretischen Annahmen des IIRM (z.B. Kessels & Hannover, 2006; Kessels et al., 2014) und der IBMT (z.B. Elmore & Oyserman, 2012; Oyserman & Destin, 2010) mit ungünstigeren Bildungsergebnissen für Jungen als Mädchen einhergeht. Basierend auf den theoretischen Annahmen sollte dies besonders bei maskulinen Jungen der Fall sein und schlechtere Noten im ebenfalls feminin wahrgenommenen Fach Deutsch als im maskulin wahrgenommenen Fach Mathematik vorhersagen. Um die Passung zwischen potentiell femininer Schule und schulischem Lernen auf der einen Seite und der Identität männlicher Jugendlicher auf der anderen Seite differenzierter analysieren zu können, wurde Maskulinität in den zwei Subskalen sozial erwünschte und sozial unerwünschte Maskulinität im Selbstbericht erhoben. Besonders Letztere sollte mit einer feminin stereotypisierten Schule im Konflikt stehen.

Die dritte Studie (*Do teachers equate male and masculine with lower academic engagement? How student's gender enactment triggers gender stereotypes at school*) untersucht, ob Gender Enactment von Jugendlichen geschlechtsstereotype Erwartungen bei Lehrkräften bezüglich des Lernverhaltens von Jungen und Mädchen hervorruft. Lehrereinschätzungen des Lernverhaltens fließen mittelbar in die Beurteilung der Schülerinnen und Schüler ein und tragen so zu Unterschieden im Schulerfolg bei (z.B. Downey & Vogt Yuan, 2005). Ausgehend vom Kontinuum-Modell der Eindrucksbildung (Fiske & Neuberg, 1990) wurde in einer experimentellen Vignettenstudie mit  $N = 104$  Gymnasiallehrkräften geprüft, ob die Eindrucksbildung von Lehrkräfte stärker auf Geschlechterstereotypen basiert, wenn Jungen und Mädchen ihre Geschlechtszugehörigkeit

im Schulkontext inszenierten, als wenn nicht. Da Stereotype männliche Schüler als wenig engagiert und störend beschreiben (z.B. Hartley & Sutton, 2013; Latsch & Hannover, 2014), sollten die Lehrkräfte – bei Aktivierung des Stereotyps – Jungen das schlechteste Lernverhalten zuschreiben. Die Studie hat ein  $2 \times 2 \times 2 \times 2$ -faktorielles Design mit den Zwischensubjektfaktoren Geschlecht Targetperson, Gender Enactment (ja/nein) und Geschlecht Lehrkraft sowie dem zugeschriebenen lernförderlichen und lernhinderlichen Verhalten als Innersubjektfaktor. Das Stimulusmaterial und die verwendeten abhängigen Variablen wurden an drei Stichproben bestehend aus Lehramtsstudierenden vorgetestet ( $N = 104$ ;  $N = 82$ ;  $N = 86$ ).



## Literaturverzeichnis I

- Adler, P. A., Kless, S. J. & Adler, P. (1992). Socialization to gender roles: Popularity among elementary school boys and girls. *Sociology of Education*, 65, 169.
- Akrami, N. & Ekehammar, B. (2005). The association between implicit and explicit prejudice: the moderating role of motivation to control prejudiced reactions. *Scandinavian journal of psychology*, 46, 361–366.
- Ambady, N., Shih, M., Kim, A. & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12, 385–390.
- Archambault, I., Janosz, M., Fallu, J.-S. & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, 32, 651–670.
- Asendorpf, J. B. (2015). Persönlichkeit. In M. A. Wirtz (Hrsg.), *Dorsch-Lexikon der Psychologie* (S. 1170). Bern: Verlag Hans Huber.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64, 359–372.
- Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
- Bahrdt, H. P. (1984). *Schlüsselbegriffe der Soziologie. Eine Einführung mit Lehrbeispielen*. München: Beck.
- Battle, E. S. (1966). Motivational determinants of academic competence. *Journal of Personality and Social Psychology*, 4, 634–642.
- Baudson, T. G. & Preckel, F. (2013). Teachers' implicit personality theories about the gifted: An experimental approach. *School Psychology Quarterly*, 28, 37–46.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006a). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Stanat, P. & Watermann, R. (2006b). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Beaman, R., Wheldall, K. & Kemp, C. (2006). Differential teacher attention to boys and girls in the classroom. *Educational Review*, 58, 339–366.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. *Zeitschrift für Pädagogische Psychologie*, 20, 233–242.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104, 682–699.
- Bem, S. L. (1983). Gender Schema Theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs*, 8, 598–616.
- Bennett, R. E., Gottesman, R. L., Rock, D. A. & Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. *Journal of Educational Psychology*, 85, 347–356.

- Berndt, T. J. & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology*, 82, 319–326.
- Blanton, H., Jaccard, J., Gonzales, P. M. & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, 42, 192–212.
- Bloom, H. S., Hill, C. J., Black, A. R. & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1, 289–328.
- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D., Prenzel, M. et al. (2009). *Geschlechterdifferenzen im Bildungssystem. Jahresgutachten 2009*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- BMBF. (2015). *Für mehr Bildungsgerechtigkeit in Deutschland*. Zugriff am 05.04.2015. Verfügbar unter <http://www.bmbf.de/de/22887.php>
- Böhme, K., Richter, D., Weirich, S., Haag, N., Wendt, H., Bos, W. et al. (2014). Messen wir dasselbe? Zur Vergleichbarkeit des IQB-Ländervergleichs 2011 mit den internationalen Studien IGLU und TIMSS 2011. *Unterrichtswissenschaft*, 42, 342–365.
- Böhme, K. & Roppelt, A. (2012). Geschlechtsbezogene Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 173–190). Münster: Waxmann.
- Böhnel, E. (1996). Die Frage der Prognostizierbarkeit von Schulerfolg in der Sekundarstufe I aufgrund der Benotung in der Primarstufe. *Unterrichtswissenschaft*, 24, 343–360.
- Bollen, K. A. & Pearl, J. (2013). Eight myths about causality and Structural Equation Models. In S. L. Morgan (Hrsg.), *Handbook of causal analysis for social research* (S. 301–328). Dordrecht: Springer Netherlands.
- Bornholt, L. & Möller, J. (2003). Attributions about achievement and intentions about further study in social context. *Social Psychology of Education*, 6, 217–231.
- Bos, W., Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 91–136). Münster: Waxmann.
- Bowen, W. G., Chingos, M. M. & McPherson, M. S. (2011). Test scores and high school grades as predictors. In W. G. Bowen, M. M. Chingos & M. S. McPherson (Hrsg.), *Crossing the finish line. Completing college at America's public universities* (S. 112–133). Princeton, NJ: Princeton University Press.
- Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual review of psychology*, 53, 371–399.
- Brehl, T., Wendt, H. & Bos, W. (2012). Geschlechtsspezifische Unterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In W. Bos, H. Wendt, O. Köller & C. Selzer (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 203–230). Münster: Waxmann.

- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik- die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, Heuvel-Panhuizen, Marja van den, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 219–249). Weinheim: Beltz.
- Brewer, M. B. (1988). A dual model of impression formation. In T. K. Srull & R. S. Wyer (Hrsg.), *Advances in social cognition* (S. 1–36). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30, 123–142.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7, 279–301.
- Brophy, J. E. (1985). Interactions of male and female students with male and female teachers. In L. C. Wilkinson & C. B. Marrett (Hrsg.), *Gender influences in classroom interaction* (S. 115–142). Orlando, FL: Academic Press.
- Brunner, M., Gogol, K. M., Sonnleitner, P., Keller, U., Krauss, S. & Preckel, F. (2013). Gender differences in the mean level, variability, and profile shape of student achievement: Results from 41 countries. *Intelligence*, 41, 378–395.
- Brunner, M., Krauss, S. & Kunter, M. (2008). Gender differences in mathematics: Does the story need to be rewritten? *Intelligence*, 36, 403–421.
- Calvin, C. M., Fernandes, C., Smith, P., Visscher, P. M. & Deary, I. J. (2010). Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. *Intelligence*, 38, 424–432.
- 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.
- Chamorro-Premuzic, T. & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality*, 17, 237–250.
- 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.
- Cho, I. (2012). The effect of teacher–student gender matching: Evidence from OECD countries. *Economics of Education Review*, 31, 54–67.
- Cizek, G. J., Fitzgerald, S. M. & Rachor, R. A. (1995). Teachers' Assessment Practices: Preparation, Isolation, and the Kitchen Sink. *Educational Assessment*, 3, 159–179.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Condry, J. & Condry, S. (1976). Sex differences: A study of the eye of the beholder. *Child Development*, 47, 812–819.
- Costa, P., Terracciano, A. & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322–331.
- Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S. & Snowden, R. J. (2010). Faking of the Implicit Association Test is statistically detectable and partly correctable. *Basic and Applied Social Psychology*, 32, 302–314.

- Cvencek, D., Meltzoff, A. N. & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development*, 82, 766–779.
- De Fruyt, F., van Leeuwen, K., Bolle, M. de & Clercq, B. de. (2008). Sex differences in school performance as a function of conscientiousness, imagination and the mediating role of problem behaviour. *European Journal of Personality*, 22, 167–184.
- de Zeeuw, E. L., van Beijsterveldt, C. E., Glasner, T. J., Bartels, M., de Geus, E. J. & Boomsma, D. I. (2014). Do children perform and behave better at school when taught by same-gender teachers? *Learning and Individual Differences*, 36, 152–156.
- Deary, I. J., Strand, S., Smith, P. & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 13–21.
- Deaux, K. & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review*, 94, 369–389.
- Deci, E. L. & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19, 109–134.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2003). Wahrgenommene Lehrereinschätzungen und das Fähigkeitsselbstkonzept von Jungen und Mädchen in der Grundschule. *Psychologie in Erziehung und Unterricht*, 50, 182–190.
- Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41, 417–440.
- Dipboye, R. L., Fromkin, H. L. & Wiback, K. (1975). Relative importance of applicant sex, attractiveness, and scholastic standing in evaluation of job applicant resumes. *Journal of Applied Psychology*, 60, 39–43.
- DiPrete, T. A. & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, 41, 1–15.
- Ditton, H. (2007). Schulübertritte, Geschlecht und soziale Herkunft. In H. Ditton (Hrsg.), *Kompetenzaufbau und Laufbahnen im Schulsystem. Ergebnisse einer Längsschnittuntersuchung an Grundschulen* (S. 63–87). Münster: Waxmann.
- Downey, D. B. & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly*, 46, 299–321.
- Dresel, M., Stoeger, H. & Ziegler, A. (2006). Klassen- und Schulunterschiede im Ausmaß von Geschlechterdiskrepanzen bei Leistungsbewertungen und Leistungsaspirationen: Ergebnisse einer Mehrebenenanalyse. *Psychologie in Erziehung und Unterricht*, 53, 44–61.
- Dresel, M. & Ziegler, A. (2006). Langfristige Förderung von Fähigkeitsselbstkonzept und impliziter Fähigkeitstheorie durch computerbasiertes attributionales Feedback. *Zeitschrift für Pädagogische Psychologie*, 20, 49–63.
- Driessen, G. (2007). The feminization of primary education: Effects of teachers' sex on pupil achievement, attitudes and behaviour. *International Review of Education*, 53, 183–203.
- Duckworth, A. L., Quinn, P. D. & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, 104, 439–451.
- Duckworth, A. L. & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16, 939–944.

- Duckworth, A. L. & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98, 198–208.
- Dumont, H., Trautwein, U., Lüdtke, O., Neumann, M., Niggli, A. & Schnyder, I. (2012). Does parental homework involvement mediate the relationship between family background and educational outcomes? *Contemporary Educational Psychology*, 37, 55–69.
- Durik, A. M., Vida, M. & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology*, 98, 382–393.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040–1048.
- Eccles, J. S. (2007). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Hrsg.), *Handbook of competence and motivation* (S. 105–121). New York, NY: Guilford Press.
- Eccles, J. S. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, 35, 195–201.
- Eccles, J. S., Adler, T. & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46, 26–43.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives. Psychological and sociological approaches* (S. 76–146). San Francisco, CA: W.H. Freeman.
- Eccles, J. S. & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, 21, 215–225.
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.
- Eckes, T. & Trautner, H. M. (2000a). Developmental social psychology of gender: An integrative framework. In T. Eckes & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 3–32). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Eckes, T. & Trautner, H. M. (Hrsg.). (2000b). *The developmental social psychology of gender*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Egan, S. K. & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, 37, 451–463.
- Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft*, 8, 521–539.
- Ehrenberg, R. G., Goldhaber, D. D. & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the national educational longitudinal study of 1988. *Industrial and Labor Relations Review*, 48, 547–561.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34, 169–189.

- Elmore, K. C. & Oyserman, D. (2012). If 'we' can succeed, 'I' can too: Identity-based motivation and gender in the classroom. *Contemporary Educational Psychology*, 37, 176–185.
- Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103–127.
- Epstein, D. (1998). Real boys don't work: 'underachievement', masculinity and the harassment of 'sissies'. In D. Epstein, J. Elwood, V. Hey & J. Maw (Hrsg.), *Failing boys? Issues in gender and achievement* (S. 96–108). Buckingham: Open University Press.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. New York, NY: Norton.
- Erikson, E. H. (1988). *Der vollständige Lebenszyklus*. Frankfurt am Main: Suhrkamp.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429–456.
- Fiedler, K. & Bless, H. (2003). Soziale Kognition. In W. Stroebe, K. Jonas & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (4. Aufl., S. 125–164). Heidelberg: Springer.
- Fiedler, K., Messner, C. & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, 17, 74–147.
- Finkenauer, C., Engels, R. C., Meeus, W. & Oosterwegel, A. (2002). Self and identity in early adolescence: The pains and gains of knowing who and what you are. In T. M. Brinthaupt & R. P. Lipka (Hrsg.), *Understanding early adolescent self and identity. Applications and interventions* (S. 25–56). Albany, NY: State University of New York Press.
- Fiske, S. T., Lin, M. & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Hrsg.), *Dual-process theories in social psychology* (S. 231–254). New York, NY: Guilford Press.
- Fiske, S. T. & Neuberg, S. L. (1990). A continuum of impression formation, from category based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74.
- Francis, B. (1999). Lads, lasses and (new) labour: 14-16-year-old students' responses to the 'laddish behaviour and boys' underachievement' debate. *British Journal of Sociology of Education*, 20, 355–371.
- Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109.
- Fredricks, J. A. & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 38, 519–533.
- Freudenthaler, H. H., Spinath, B. & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality*, 22, 231–245.
- Freund, P. A., Holling, H. & Preckel, F. (2007). A multivariate, multilevel analysis of the relationship between cognitive abilities and scholastic achievement. *Journal of Individual Differences*, 28, 188–197.
- Frey, M. C. & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, 15, 373–378.

- Frome, P. M. & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74, 435–452.
- Fuchs, G. (2014, September). *Psychometrische Datenqualität bildungsstandardbasierter Testverfahren*. Vortrag auf der 1. Tagung des Berlin Interdisciplinary Education Research Network, Berlin.
- Goffman, E. (1976). *Gender advertisements*. New York, NY: Harper & Row.
- Gold, D. & Reis, M. (1982). Male teacher effects on young children: A theoretical and empirical consideration. *Sex Roles*, 8, 493–513.
- Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24, 13–23.
- Greene, B. A., DeBacker, T. K., Ravindran, B. & Krows, A. J. (1999). Goals, values, and beliefs as predictors of achievement and effort in high school mathematics classes. *Sex Roles*, 40, 421–458.
- Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L. & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97, 17–41.
- Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4. Aufl.). New York, NY: Psychology Press.
- Hannover, B. & Kessels, U. (2002). Challenge the science-stereotype! Der Einfluss von Technik-Freizeitkursen auf das Naturwissenschaften-Stereotyp von Schülerinnen und Schülern. *Zeitschrift für Pädagogik - Beiheft*, 45, 341–358.
- Hannover, B. & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction*, 14, 51–67.
- 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.
- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L. & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, 100, 105–122.
- Hartley, B. L. & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development*, 84, 1716–1733.
- Havighurst, R. J. (1948). *Developmental tasks and education*. Chicago, IL: University of Chicago Press.
- Havighurst, R. J. (1956). Research on the developmental-task concept. *The School Review*, 64, 215–223.

- Helbig, M. (2010a). Geschlecht der Lehrer und Kompetenzentwicklung der Schüler. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 273–288). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Helbig, M. (2010b). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 93–111.
- Helbig, M. (2012). Boys do not benefit from male teachers in their reading and mathematics skills: Empirical evidence from 21 European Union and OECD countries. *British Journal of Sociology of Education*, 33, 661–677.
- Herbert, J. & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, 26, 276–295.
- Hill, C. J., Bloom, H. S., Black, A. R. & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2, 172–177.
- Hill, J. & Lynch, M. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Petersen (Hrsg.), *Girls at puberty. Biological and psychosocial perspectives* (S. 201–228). New York, NY: Plenum.
- Hilliard, L. J. & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, 81, 1787–1798.
- Hilton, J. L. & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47, 237–271.
- Hofer, M., Kuhnle, C., Kilian, B. & Fries, S. (2012). Cognitive ability and personality variables as predictors of school grades and test scores in adolescents. *Learning and Instruction*, 22, 368–375.
- Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 217–244). Münster: Waxmann.
- Hosenfeld, I., Köller, O. & Baumert, J. (1999). Why sex differences in mathematics achievement disappear in German secondary schools: A reanalysis of the German TIMSS-data. *Studies in Educational Evaluation*, 25, 143–161.
- Hoyle, R. H. (1995). The Structural Equation Modeling Approach. Basic concepts and fundamental issues. In R. H. Hoyle (Hrsg.), *Structural equation modeling. Concepts, issues, and applications* (S. 1–36). Thousand Oaks, CA: Sage Publications.
- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, 49, 505–528.
- Hülür, G., Wilhelm, O. & Robitzsch, A. (2011). Multivariate Veränderungsmodelle für Schulnoten und Schülerleistungen in Deutsch und Mathematik. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 173–185.
- Huyge, E., van Maele, D. & van Houtte, M. (2014). Does students' machismo fit in school? Clarifying the implications of traditional gender role ideology for school belonging. *Gender and Education*, 1–18.
- Hyde, J. S. (2005). The Gender Similarities Hypothesis. *American Psychologist*, 60, 581–592.



- Hyde, J. S. & Kling, K. C. (2001). Women, motivation, and achievement. *Psychology of Women Quarterly*, 25, 364–378.
- Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik* (6. Aufl.). Weinheim: Beltz.
- Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy. *Gender and Education*, 14, 37–50.
- Jackson, C. (2003). Motives for 'laddishness' at school: Fear of failure and fear of the 'feminine'. *British Educational Research Journal*, 29, 583–598.
- Jackson, C. & Dempster, S. (2009). 'I sat back on my computer ... with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, 18, 341–356.
- Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, 83, 518–527.
- Jacobs, J. E. & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, 63, 932–944.
- Jensen, A. R. (1998). *The g factor. The science of mental ability*. Westport, CT: Praeger.
- Jewell, J. A. & Brown, C. S. (2014). Relations among gender typicality, peer relations, and mental health during early adolescence. *Social Development*, 23, 137–156.
- 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.
- Jussim, L. & Eccles, J. S. (1992). Teacher expectations: II. Construction and reflection of student achievement. *Journal of Personality and Social Psychology*, 63, 947–961.
- Jussim, L., Eccles, J. S. & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations. Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 28, 281–388.
- Jussim, L. & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131–155.
- Kaiser, J., Retelsdorf, J., Südkamp, A. & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, 28, 73–84.
- Kedar-Voivodas, G. (1983). The impact of elementary children's school roles and sex roles on teacher attitudes: An interactional analysis. *Review of Educational Research*, 53, 415–437.
- Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M. & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, 42, 11–26.
- Kerger, S., Martin, R. & Brunner, M. (2011). How can we enhance girls' interest in scientific topics? *British Journal of Educational Psychology*, 81, 606–628.
- 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. & Hannover, B. (2004). Entwicklung schulischer Interessen als Identitätsregulation. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (S. 398–412). Münster: Waxmann.
- Kessels, U. & Hannover, B. (2006). Zum Einfluss des Image von mathematisch-naturwissenschaftlichen Schulfächern auf die schulische Interessensentwicklung. In M. Prenzel (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 350–369). Münster: Waxmann.
- Kessels, U. & Hannover, B. (2007). How the image of math and science affects the development of academic interest. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report of the DFG priority programme* (S. 283–297). Münster: Waxmann.
- Kessels, U., Heyder, A., Latsch, M. & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, 56, 219–228.
- Kessels, U., Rau, M. & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 76, 761–780.
- 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.
- Kessels, U., Warner, L. M., Holle, J. & Hannover, B. (2008). Identitätsbedrohung durch positives schulisches Leistungsfeedback. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40, 22–31.
- Kiefer, A. K. & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes. A prospective study of female college students. *Psychological Science*, 18, 13–18.
- Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, 105, 198–214.
- Klauer, K. J. & Leutner, D. (2010). Intelligenz und Begabung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 304–310). Weinheim: Beltz.
- Kleickmann, T., Brehl, T., Saß, S., Prenzel, M. & Köller, O. (2012). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selzer (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 123–170). Münster: Waxmann.
- Kling, K. C., Nofhle, E. E. & Robins, R. W. (2013). Why do standardized tests underpredict women's academic performance? The role of conscientiousness. *Social Psychological and Personality Science*, 4, 600–606.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Hrsg.), *The development of sex differences* (S. 82–172). Stanford, CA: Stanford University Press.
- Köller, O., Baumert, J. & Bos, W. (2002). TIMSS- Third International Mathematics and Science Study. Dritte internationale Mathematik- und Naturwissenschaftsstudie. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Aufl., S. 269–284). Weinheim: Beltz.

- Köller, O., Daniels, Z., Schnabel, K. U. & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 14, 26–37.
- Köller, O., Knigge, M. & Tesch, B. *Sprachliche Kompetenzen im Ländervergleich. Befunde des ersten Ländervergleichs zur Überprüfung der Bildungsstandards für den Mittleren Schulabschluss in den Fächern Deutsch, Englisch und Französisch. Zusammenfassung.* Verfügbar unter [https://www.iqb.hu-berlin.de/laendervergleich/LV08\\_09/LV\\_ZF\\_0809c.pdf](https://www.iqb.hu-berlin.de/laendervergleich/LV08_09/LV_ZF_0809c.pdf)
- Krahé, B., Berger, A. & Möller, I. (2007). Entwicklung und Validierung eines Inventars zur Erfassung des Geschlechtsrollen-Selbstkonzepts im Jugendalter. *Zeitschrift für Sozialpsychologie*, 38, 195–208.
- Krapp, A. (2010). Interesse. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 311–323). Weinheim: Beltz.
- Kuhl, P. & Hannover, B. (2012). Differenzielle Benotungen von Mädchen und Jungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 153–162.
- Lam, S.-f., Jimerson, S., Kikas, E., Cefai, C., Veiga, F. H., Nelson, B. et al. (2012). Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of School Psychology*, 50, 77–94.
- Lam, Y. H. R., Tse, S. K., Lam, J. W. I. & Loh, E. K. Y. (2010). Does the gender of the teacher matter in the teaching of reading literacy? Teacher gender and pupil attainment in reading literacy in Hong Kong. *Teaching and Teacher Education*, 26, 754–759.
- Lane, K. A., Goh, J. X. & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles*, 66, 220–234.
- 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.
- LeBel, E. P. & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin*, 37, 570–583.
- Lehmann, R. H., Peek, R. & Gänsfuß, R. (1997). *Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen.* Hamburg: Behörde für Schule, Jugend und Berufsbildung.
- Lightbody, P., Siann, G., Stocks, R. & Walsh, D. (1996). Motivation and attribution at secondary school: The role of gender. *Educational Studies*, 22, 13–25.
- Lüdtke, O., Köller, O., Artelt, C., Stanat, P. & Baumert, J. (2002). Eine Überprüfung von Modellen zur Genese akademischer Selbstkonzepte: Ergebnisse aus der PISA-Studie. *Zeitschrift für Pädagogische Psychologie*, 16, 151–164.
- Maaz, K., Baeriswyl, F. & Trautwein, U. (2011). *Herkunft zensiert? Leistungsdiagnostik und soziale Ungleichheiten in der Schule Eine Studie im Auftrag der Vodafone Stiftung Deutschland.* Verfügbar unter <http://www.vodafone-stiftung.de/publikationmodul/detail/33.html>

- Maaz, K. & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems. Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In Bundesministerium für Bildung und Forschung (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 151–180). Berlin: Bundesministerium für Bildung und Forschung.
- Maaz, K., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2, 99–106.
- MacKinnon, D. P., Krull, J. L. & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, 1, 173–181.
- Madon, S., Willard, J., Guyll, M. & Scherr, K. C. (2011). Self-Fulfilling Prophecies: Mechanisms, Power, and Links to Social Problems. *Social and Personality Psychology Compass*, 5, 578–590.
- Markus, H. & Nurius, P. (1986). Possible selves. *American Psychologist*, 41, 954–969.
- Marsh, H. W., Martin, A. J. & Cheng, J. H. S. (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.
- Marsh, H. W., Nagengast, B. & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, 49, 1194–1218.
- Martin, C. L. & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, 52, 1119–1134.
- Martino, W. J. (2008). Male teachers as role models: Addressing issues of masculinity, pedagogy and the re-masculinization of schooling. *Curriculum Inquiry*, 38, 189–223.
- Matthews, J. S., Ponitz, C. C. & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, 101, 689–704.
- McClure, J., Meyer, L. H., Garisch, J., Fischer, R., Weir, K. F. & Walkey, F. H. (2011). Students' attributions for their best and worst marks: Do they relate to achievement? *Contemporary Educational Psychology*, 36, 71–81.
- McCrae, R. R. & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81–90.
- McDermott, P. A., Mordell, M. & Stoltzfus, J. C. (2001). The organization of student performance in American schools: Discipline, motivation, verbal learning, nonverbal learning. *Journal of Educational Psychology*, 93, 65–76.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20, 20–32.
- McMillan, J. H., Myran, S. & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95, 203–213.
- Meece, J. L., Parsons, J. E., Kaczala, C. M. & Goff, S. B. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin*, 91, 324–348.
- Meece, J. L., Wigfield, A. & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, 82, 60–70.

- Mickelson, R. A. (1989). Why does Jane read and write so well? The anomaly of women's achievement. *Sociology of Education*, 62, 47–63.
- Ministerium für Schule und Weiterbildung NRW. (2013). *Zentralabitur an Gymnasien und Gesamtschulen. Ergebnisse 2013*. Verfügbar unter <http://www.standardsicherung.schulministerium.nrw.de/abitur/upload/download/Zentralabitur-Gymnasiale-Oberstufe-2013.pdf>
- Möller, J., Pohlmann, B., Köller, O. & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79, 1129–1167.
- Müller, K. & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 245–274). Münster: Waxmann.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O. & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12, 323–345.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77–101.
- Neuenschwander, M. P. & Malti, T. (2009). Selektionsprozesse beim Übergang in die Sekundarstufe I und II. *Zeitschrift für Erziehungswissenschaft*, 12, 216–232.
- 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.
- Neugebauer, M., Helbig, M. & Landmann, A. (2011). Unmasking the myth of the same-sex teacher advantage. *European Sociological Review*, 27, 669–689.
- Neumann, M., Milek, A., Maaz, K. & Gresch, C. (2010). Zum Einfluss der Klassenzusammensetzung auf den Übergang von der Grundschule in die weiterführenden Schulen. In Bundesministerium für Bildung und Forschung (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 229–252). Berlin: Bundesministerium für Bildung und Forschung.
- Nicholls, J. G., Patashnick, M. & Nolen, S. B. (1985). Adolescents' theories of education. *Journal of Educational Psychology*, 77, 683–692.
- Nie, Y. & Liem, G. A. D. (2013). Extending antecedents of achievement goals: The double-edged sword effect of social-oriented achievement motive and gender differences. *Learning and Individual Differences*, 23, 249–255.
- Niklas, F. & Schneider, W. (2012). Die Anfänge geschlechtsspezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 123–138.
- Nosek, B. A. & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition*, 19, 625–666.
- Nosek, B. A., Banaji, M. R. & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, 83, 44–59.
- Nosek, B. A. & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal*, 48, 1125–1156.

- OECD. (2013b). *PISA 2012 results. Excellence through equity (Volume II)*. Paris: OECD Publishing.
- OECD. (2013a). *PISA 2012 results: What students know and can do (Volume I)*. Paris: OECD Publishing.
- OECD. (2010). *PISA 2009 results: What students know and can do- student performance in reading, mathematics and science (Volume I)*. Paris: OECD Publishing.
- Orr, A. J. (2011). Gendered capital: Childhood socialization and the “boy crisis” in education. *Sex Roles*, 65, 271–284.
- Ostendorf, F. & Angleitner, A. (2004). NEO-Personality Inventory Revised. *NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung (NEO-PI-R)*. Göttingen: Hogrefe.
- Oyserman, D. (2009a). Identity-based motivation and consumer behavior. *Journal of Consumer Psychology*, 19, 276–279.
- Oyserman, D. (2009b). Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior. *Journal of Consumer Psychology*, 19, 250–260.
- Oyserman, D. & Destin, M. (2010). Identity-Based Motivation: Implications for intervention. *The Counseling Psychologist*, 38, 1001–1043.
- Oyserman, D., Fryberg, S. A. & Yoder, N. (2007). Identity-based motivation and health. *Journal of Personality and Social Psychology*, 93, 1011–1027.
- Passolunghi, M. C., Rueda Ferreira, T. I. & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences*, 34, 70–76.
- Plante, I., de la Sablonnière, R., Aronson, J. M. & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology*, 38, 225–235.
- Plante, I., O’Keefe, P. A. & Théorêt, M. (2013). The relation between achievement goal and expectancy-value theories in predicting achievement-related outcomes: A test of four theoretical conceptions. *Motivation and Emotion*, 37, 65–78.
- Platt, C. W. (1988). Effects of causal attributions for success on first-term college performance: A covariance structure model. *Journal of Educational Psychology*, 80, 569–578.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135, 322–338.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M. et al. (Hrsg.). (2006). *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.
- Randall, J. & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26, 1372–1380.
- Räty, H., Vänskä, J., Kasanen, K. & Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: A replication and extension of Yee and Eccles. *Sex Roles*, 46, 121–128.

- Reilly, D., Neumann, D. L. & Andrews, G. (2014). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology*, Advance online publication.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, Heuvel-Panhuizen, Marja van den, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 120–141). Weinheim: Beltz.
- Retelsdorf, J., Schwartz, K. & Asbrock, F. (2015). “Michael can’t read!” Teachers’ gender stereotypes and boys’ reading self-concept. *Journal of Educational Psychology*, *107*, 186–194.
- Robinson, J. P. & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*, 268–302.
- Rodax, K. & Hurrelmann, K. (1986). Die Bildungsbeteiligung der Mädchen und Frauen- ein Indikator für wachsende Chancengleichheit? *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie*, *6*, 138–146.
- Sälzer, C. & Prenzel, M. (2013). PISA 2012- eine Einführung in die aktuelle Studie. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 11–46). Münster: Waxmann.
- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 47–98). Münster: Waxmann.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*, 299–323.
- Schiepe-Tiska, A., Schöps, K., Rönnebeck, S., Köller, O. & Prenzel, M. (2013). Naturwissenschaftliche Kompetenz in PISA 2012: Ergebnisse und Herausforderungen. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 189–216). Münster: Waxmann.
- Schilling, S. R., Sparfeldt, J. R. & Rost, D. H. (2006). Facetten schulischen Selbstkonzepts. *Zeitschrift für Pädagogische Psychologie*, *20*, 9–18.
- Schmitt, D. P., Realo, A., Voracek, M. & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, *94*, 168–182.
- Schneider, D. J. (2004). *The psychology of stereotyping*. Princeton, N.J.: The Guilford Press.
- Schuler, H. (2010). Noten als Prädiktoren von Studien- und Berufserfolg. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 599–606). Weinheim: Beltz.
- Schulgesetz für das Land Nordrhein-Westfalen. (2005). Schulgesetz NRW - SchulG. Verfügbar unter <http://www.schulministerium.nrw.de/docs/Recht/Schulrecht/Schulgesetz/Schulgesetz.pdf>

- Selter, C., Walther, G., Wessel, J. & Wendt, H. (2012). Mathematische Kompetenzen im internationalen Vergleich. Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 69–122). Münster: Waxmann.
- Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles*, 49, 451–463.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417–453.
- Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles*, 67, 617–629.
- Spence, J. T., Helmreich, R. L. & Holahan, C. K. (1979). Negative and positive components of psychological masculinity and femininity and their relationships to self-reports of neurotic and acting out behaviors. *Journal of Personality and Social Psychology*, 37, 1673–1682.
- Spengler, M., Lüdtke, O., Martin, R. & Brunner, M. (2013). Personality is related to educational outcomes in late adolescence: Evidence from two large-scale achievement studies. *Journal of Research in Personality*, 47, 613–625.
- Spinath, B., Eckert, C. & Steinmayr, R. (2014). Gender differences in school success: what are the roles of students' intelligence, personality and motivation? *Educational Research*, 56, 230–243.
- Spinath, B., Spinath, F. M., Harlaar, N. & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34, 363–374.
- Spinath, B., Stiensmeier-Pelster, J., Schöne, C. & Dickhäuser, O. (2002). *SELLMO- Skalen zur Erfassung der Lern- und Leistungsmotivation*. Göttingen: Hogrefe.
- Stanat, P. & Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 249–269). Opladen: Leske + Budrich.
- Statistisches Bundesamt. (1982). *Statistisches Jahrbuch 1981 für die Bundesrepublik Deutschland*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (1983). *Statistisches Jahrbuch 1982 für die Bunderepublik Deutschland*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2013a). *Bildung und Kultur. Allgemeinbildende Schulen Schuljahr 2012/2013*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2013b). *Datenreport 2013. Ein Sozialbericht für die Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung.
- Statistisches Bundesamt. (2014). *Schulen auf einen Blick*. Wiesbaden: Statistisches Bundesamt.
- Steele, J. (2003). Children's gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology*, 33, 2587–2606.



- Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, *51*, 165–179.
- Steffens, M. C. & 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. C., 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., Dinger, F. C. & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality*, *26*, 335–349.
- Steinmayr, R. & Meißner, A. (2013). Zur Bedeutung der Intelligenz und des Fähigkeitsselbstkonzeptes bei der Vorhersage von Leistungstests und Noten in Mathematik. *Zeitschrift für Pädagogische Psychologie*, *27*, 273–282.
- 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.
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*, 80–90.
- Steinmayr, R. & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica*, *56*, 195–211.
- Steinmayr, R., Wirthwein, L. & Schöne, C. (2014). Gender and numerical intelligence: Does motivation matter? *Learning and Individual Differences*, *32*, 140–147.
- Stern, E. (2001). Intelligenz, Wissen, Transfer und der Umgang mit Zeichensystemen. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S. 163–204). Lengerich: Pabst.
- Sternberg, R. J., Grigorenko, E. & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, *47*, 1–41.
- Strand, S., Deary, I. J. & Smith, P. (2006). Sex differences in cognitive abilities test scores: a UK national picture. *The British journal of educational psychology*, *76*, 463–480.
- Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*, 743–762.
- Sun, Y. & Cheng, L. (2014). Teachers' grading practices: meaning and values assigned. *Assessment in Education: Principles, Policy & Practice*, *21*, 326–343.
- Taconis, R. & Kessels, U. (2009). How choosing science depends on students' individual fit to 'science culture'. *International Journal of Science Education*, *31*, 1115–1132.
- Takei, Y., Johnson, M. P. & Clark, M. E. (1998). Academic achievement and impression management as factors in the grading of White junior high pupils. *Sociological Perspectives*, *41*, 27–48.
- Tent, L. & Birkel, P. (2010). Zensuren. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 949–958). Weinheim: Beltz.
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*, 144–151.

- Trapmann, S., Hell, B., Weigand, S. & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21, 11–27.
- Trautwein, U., Lüdtke, O., Kastens, C. & Köller, O. (2006). Effort on homework in grades 5–9: Development, motivational antecedents, and the association with effort on classwork. *Child Development*, 77, 1094–1111.
- Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I. & Niggli, A. (2009). Different forces, same consequence: conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology*, 97, 1115–1128.
- Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G. & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology*, 104, 763–777.
- Urhahne, D. (2008). Sieben Arten der Lernmotivation. *Psychologische Rundschau*, 59, 150–166.
- Valtin, R., Wagner, C. & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse- schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 187–238). Münster: Waxmann.
- van den Ham, A.-K., Nissen, A., Ehmke, T., Sälzer, C. & Roppelt, A. (2014). Mathematische Kompetenz in PISA, IQB-Ländervergleich und NEPS - Drei Studien, gleiches Konstrukt? *Unterrichtswissenschaft*, 42, 321–341.
- Verordnung über die Schularten und Bildungsgänge der Sekundarstufe I. (2010). Sekundarstufe I-Verordnung. Sek I-VO. Verfügbar unter <http://gesetze.berlin.de/jportal/portal/t/dut/page/bsbeprod.psml?action=controls.jw.PrintOrSaveDocumentContent&case=save>
- Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204.
- Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educational Research and Evaluation*, 12, 305–322.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91, 461–481.
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6, 49–78.
- Wigfield, A. & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, 30, 1–35.
- Wigfield, A. & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12, 265–310.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.

- Wiggins, J. S. & Trapnell, P. D. (1997). Personality structure: The return of the Big Five. In R. Hogan, J. Johnson & S. R. Briggs (Hrsg.), *Handbook of personality psychology* (S. 737–765). San Diego, CA: Academic Press.
- Wilgenbusch, T. & Merrell, K. W. (1999). Gender differences in self-concept among children and adolescents: A meta-analysis of multidimensional studies. *School Psychology Quarterly, 14*, 101–120.
- Willingham, W. W., Pollack, J. M. & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement, 39*, 1–37.
- Wilson, T. D., Lindsey, S. & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
- Wirthwein, L., Sparfeldt, J. R., Piquart, M., Wegerer, J. & Steinmayr, R. (2013). Achievement goals and academic achievement: A closer look at moderating factors. *Educational Research Review, 10*, 66–89.
- Younger, M., Warrington, M. & Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education, 20*, 325–341.
- Yunger, J. L., Carver, P. R. & Perry, D. G. (2004). Does gender identity influence children's psychological well-being? *Developmental Psychology, 40*, 572–582.
- Ziegenspeck, J. W. (1999). *Handbuch Zensur und Zeugnis in der Schule*. Bad Heilbrunn: Klinkhardt.



## 5 Studien

### 5.1 Studie 1: Beyond expectancies and values: Applying the Eccles et al. model to the underachievement of boys

Note: This is the first author's version a work that has been submitted for publication.

Heyder, A., Kessels, U., & Steinmayr, R. (2014). Beyond expectancies and values: Applying the Eccles et al. model to the underachievement of boys. *Manuscript submitted for publication.*

### **Abstract**

This study tested whether the Eccles et al. expectancy-value model of achievement-related choices and performance (e.g., Eccles et al., 1983) could explain gender differences in both language and math achievement. Students' ability self-concept and task values in math and German, intelligence and prior achievement, parents' education and perception of the child's competence in math and German were assessed in  $N = 520$  German students (age  $M = 17$  years). These variables were included as multiple mediators in two structural equation models specified for explaining boys' lower grades in German and girls' lower grades in math. Boys' lower grades in German were not explained by the mediators. In math a suppression effect was found: After controlling for motivational, cognitive, and background variables, boys earned lower grades than girls not only in German but also in math. Future research on boys' underachievement should therefore focus more on domain-unspecific motivational factors influencing achievement.

Keywords: gender differences, academic achievement, expectancy-value model, intelligence, underachievement of boys

**Highlights:**

- Boys' lower grades in German are not explained by the Eccles et al. model.
- Girls' lower grades in math are over-explained by the Eccles et al. model.
- Boys' academic underachievement seems to be due to domain-unspecific factors.

## 1. Introduction

The study of gender differences in academic outcomes has a long history in educational research. For many decades, researchers have focused on girls' underperformance and underrepresentation in math. Taken together, findings show that in many countries boys outperform girls on standardized math achievement tests (e.g., PISA 2012, OECD, 2013a), score higher on math ability self-concept and task values (e.g., Nagy, Trautwein, Baumert, Köller, & Garrett, 2006), and report more interest in math (e.g., OECD, 2013b) as well as in pursuing math-related careers (e.g., Watt, 2006). Math is further stereotyped as a male domain, and males are considered to be more talented in math than females (e.g., Steffens, Jelenec, & Noack, 2010). Although boys indeed demonstrate greater competence on standardized math tests in many countries, they do not consistently earn higher grades than girls (see e.g., Kenney-Benson, Pomerantz, Ryan, & Patrick, 2006; Kimball, 1989). A recent meta-analysis on gender differences in academic achievement found that the nationality of the sample was a significant moderator of the effect of gender on math grades: Whereas girls were found to earn slightly better grades in math in North American samples, no gender differences were found on average for students from other countries (Voyer & Voyer, 2014). In Germany, some individual studies have reported small gender differences in math grades in favor of boys (e.g., Dresel, Stöger, & Ziegler, 2006; Dumont et al., 2012; Goetz, Frenzel, Hall, & Pekrun, 2008; Steinmayr & Spinath, 2008). However, even in the male-stereotyped domain of math, boys' cognitive potential seems to manifest in high grades less often than girls'.

The current discussion about boys as academic underachievers has brought a greater focus on gender differences that occur at the expense of the boys (for an overview, see Buchmann, DiPrete, & McDaniel, 2008). For example, boys graduate with lower GPAs (NCES, 2004), are underrepresented in the academic track of tracked school systems (e.g., for Germany: Statistisches Bundesamt, 2012), are more often held back a grade, and more often



drop out of school than girls (e.g., NCES, 2009). Gender differences favoring girls are especially pronounced and consistently found in the female-stereotyped domain of languages (Steffens et al., 2010), which are considered to be the basis for high achievement in any subject (e.g., OECD, 2002). Boys perform worse than girls on standardized verbal achievement tests, for example, PIRLS 2011 (Mullis, Martin, Foy, & Drucker, 2012) or PISA 2012 (OECD, 2013a); boys have a lower verbal ability self-concept and value the verbal domain less than girls do (e.g., Durik, Vida, & Eccles, 2006; Steinmayr & Spinath, 2008); and boys also earn consistently lower grades in languages than girls (e.g., Durik et al., 2006; Spinath, Freudenthaler, & Neubauer, 2010; Voyer & Voyer, 2014). These gender differences in language achievement are consistently larger than those found in math achievement (e.g., OECD, 2013a).

Gender differences in academic achievement have been successfully explained in the past by the Eccles et al. expectancy-value model of achievement-related choices and performance (in the following, called the Eccles et al. model; e.g., Eccles et al., 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). This model theoretically elaborates the way in which an individual's characteristics (e.g., motivational beliefs or aptitude) interact with socioeconomic factors and jointly contribute to gender differences in academic outcomes. Most of these studies have focused on girls' disadvantage in math. Whether boys' lower achievement in languages is due to the same psychological mechanisms is less substantiated to date, as empirical studies in this domain are scarce. Taking into account the ongoing discussion on the underachievement of boys in school, it seems fruitful to additionally test whether the Eccles et al. model can be used to explain the prevailing and substantial disadvantage of boys in the verbal domain.

Studies successfully predicting achievement by using the Eccles et al. model have focused primarily on the more proximal factors such as expectancy of success and values (e.g., Eccles, Adler, & Meece, 1984; Nagy et al., 2006; Watt, 2006). The more distal factors

included in the theoretical model have received relatively less attention in studies from this strand of research. Due to its complexity, it is impossible to test the entire Eccles et al. model at once, but it seems fruitful to extend the focus to factors that other strands of research have consistently found to be important predictors of an individual's achievement, responsible for many of the disparities in educational outcomes. Several studies that did not apply the Eccles et al. model have highlighted, for instance, the significance of a child's socioeconomic background (for a meta-analysis, see e.g., Sirin, 2005) and general cognitive abilities (e.g., Deary, Strand, Smith, & Fernandes, 2007) for academic success. In a nutshell, the goal of our study was to explain gender differences in both the mathematical and verbal domains by applying the Eccles et al. model while taking into account not only proximal factors such as expectancy of success and task values but also adding the more distal factors such as socioeconomic status (SES) and general intelligence to the model.

## **2. Explaining Gender Differences in Academic Outcomes: The Eccles et al. Model**

A great deal of research has applied the Eccles et al. model to explain gender differences in academic outcomes. Building on Atkinson's expectancy-value theory (e.g., Atkinson, 1964), Eccles and colleagues have traced (gender) differences in domain-specific academic outcomes most directly back to (gender) differences in students' expectations of success and the values students ascribed to a task (e.g., Eccles et al., 1983; Eccles & Wigfield, 2002). Expectancy of success is defined as the student's beliefs regarding how well he or she will perform on a future task and is often operationalized by the student's ability self-concept. Although expectancies for success and ability self-concept are not identical, as ability self-concept focuses on current ability and expectancies focus on the future, they are highly correlated (Wigfield & Eccles, 2000). Subjective task values can be broken down into four components: intrinsic value, which is the expected enjoyment of a task; attainment value, which is the personal importance of doing well on a task; utility value of a task for obtaining

external rewards or approaching long-term goals; and psychological cost, which means that the decision to engage in one activity limits access to other activities (e.g., Eccles, 2007).

Eccles and colleagues have integrated these motivational components into a greater framework of sociocultural factors and a child's prior experiences. More precisely, ability self-concept and subjective task values are presumed to be influenced by the cultural milieu of the child (e.g., gender role stereotypes and family characteristics), by beliefs and behaviors of those socializing the child, by the child's characteristics (e.g., aptitude or personality), by previous achievement-related experiences, and by the child's perceptions and interpretations of all of these (e.g., Wigfield & Eccles, 2000). These characteristics of the child, the child's family, and the child's learning environment should not have a direct impact on a child's academic choices and performance but should rather have an indirect effect mediated via the child's ability self-concept and task values. Several components of the Eccles et al. model have been fruitfully used to explain individual differences in academic outcomes in various academic domains. In attempts to explain *gender differences*, however, the Eccles et al. model has so far been used mostly within the male-stereotyped domains of math, science, and sport, whereas studies explaining gender differences in languages have been relatively scarce.

In what follows, we will summarize empirical findings regarding the way various components of the model contribute to explaining gender differences in math and language-related outcomes. First, we will summarize findings on the impact of the more distal factors in the model, namely child's aptitude, family characteristics, socializers' beliefs, and child's prior achievement experiences, which are not assumed to impact achievement directly but indirectly by the mediating, more proximal factors of the model (ability self-concept and task values). Second, we will summarize findings on the relevance of these proximal factors.

### **2.1 Aptitudes of the Child**

The child's aptitudes are included in the Eccles et al. model as one of the distal factors that impact academic outcomes. In general, aptitude, operationalized by intelligence test

scores, is commonly known as the strongest predictor of academic achievement (e.g., Deary et al., 2007; Kuncel, Hezlett, & Ones, 2004). Studies have usually reported effect sizes of around  $r = .5$  for the association between intelligence and educational achievement (e.g., Neisser et al., 1996), but associations as strong as  $r = .8$  have been found when both constructs were modeled as latent traits and school performance was operationalized by means of standardized school achievement tests (e.g., Calvin, Fernandes, Smith, Visscher, & Deary, 2010; Deary et al., 2007). However, when school achievement has been operationalized by means of grades, the associations have seldom exceeded  $r = .5$  on a latent level (e.g., Steinmayr & Spinath, 2008). Studies have found a stronger association between intelligence and math achievement than between intelligence and achievement in the verbal domain (Calvin et al., 2010; Spinath et al., 2010; Steinmayr & Spinath, 2008).

In the Eccles et al. model, intelligence is expected to influence academic outcomes indirectly via the child's ability self-concept, task values, and socialization agents. Research explicitly applying the Eccles et al. model has usually paid little attention to the role of intelligence in explaining gender differences in academic outcomes. This seems logical against the background that most studies have found gender differences in general intelligence to be negligible and not responsible for gender differences in academic achievement (e.g., Calvin et al., 2010; Halpern, 2012). In a sample of high school students preselected for the highest school track, however, Steinmayr and Spinath (2008) found that controlling for general intelligence increased rather than reduced gender differences in general school performance. This finding indicates that although there is generally no reason to expect that intelligence would explain gender differences in academic outcomes, general intelligence might play a role when certain selected samples are analyzed.

## **2.2 Family Characteristics**

Another more distal factor included in the Eccles et al. model refers to the broad term of *family characteristics*. Research following the Eccles et al. model has so far concentrated

on the effects and mediating variables of parents' gender stereotypes regarding math ability (e.g., Eccles, Jacobs & Harold, 1990; Jacobs, 1991; Tiedemann, 2000). These studies have recently been augmented by the finding that students' ability self-concept and task values mediate the effect of students' math and language gender stereotypes on math and language outcomes (Plante, de la Sablonnière, Aronson, & Théorêt, 2013). However, other indicators of a person's cultural milieu that are related to individual differences in academic outcomes have so far been ignored by researchers who have applied the Eccles et al. model. With regard to family characteristics, countless studies have found that parents' SES, operationalized via parents' level of education or profession, predicted their child's academic achievement. As an example, two meta-analyses reported mean correlations between SES and academic achievement of around  $r = .3$  (Sirin, 2005; White, 1982). Furthermore, in line with the Eccles et al. model, SES was found to be associated with children's ability self-concept and values, which are the more proximal predictors of academic achievement (Steinmayr, Dinger, & Spinath, 2012).

Parents' SES seems to be important not only for a child's academic achievement per se but also for the size of gender differences in academic achievement. Several studies have found the language achievement gender difference that favors girls to be more pronounced in disadvantaged samples (i.e., children from low-income families) than in advantaged samples (e.g., Entwisle, Alexander, & Olson, 2007; Mensah & Kiernan, 2010). Against this background, it is surprising that in attempts to understand the emergence of gender differences in math and language achievement, parents' SES has, to our knowledge, not yet been included in studies that have applied the Eccles et al. model. The present study fills this gap by exploring the role of parents' education in explaining gender differences in academic achievement within the theoretical framework of the Eccles et al. model.

### **2.3 Socializers' Perception of Child's Ability**

According to the socialization element of the Eccles et al. model, parents' perceptions of their child's ability have a major influence on the child's ability self-concept, task values, and, therefore, performance (e.g., Eccles et al., 1983; Frome & Eccles, 1998; Jacobs, Davis-Kean, Bleeker, Eccles, & Malanchuk, 2005). In support of predictions made by the model, the influence of parents' perceptions on a child's ability self-concept is as strong as or stronger than the influence of the child's previous achievement (e.g., Eccles et al., 1983; Frome & Eccles, 1998; Herbert & Stipek, 2005; Tiedemann, 2000). Furthermore, Frome and Eccles (1998) found in a longitudinal study that gender differences in children's verbal ability self-concepts were partially due to parents' gender-differentiated ability perceptions in the verbal domain in favor of girls. For math, however, no indirect effect of gender on ability self-concept mediated by parents' perceptions was found (Frome & Eccles, 1998). Importantly, existing studies on the impact of parents' perceptions have so far focused on children in elementary school, and the assumptions of the Eccles et al. model regarding the role of parents' perceptions have not yet been tested in older age groups. As domain-specific gender differences in academic achievement tend to grow as students move through school (e.g., Buchmann et al., 2008), the present study sought to analyze the role of socializers' perceptions of children's ability in accounting for gender differences in math and language achievement in a sample of adolescents attending secondary school.

### **2.4 Previous Achievement**

Several studies have examined whether a child's previous achievement can predict future achievement (e.g., Frome & Eccles, 1998; Huang, 2011; Kytälä & Björn, 2010). However, in their theoretical model, Eccles and colleagues emphasize the role of the child's own interpretation of previous achievement as it influences the child's ability self-concept, task values, and, thereby, future performance rather than the role of previous achievement by itself (e.g., Eccles et al., 1983). Therefore, prior achievement is not supposed to be directly

linked to future academic outcomes but should rather be indirectly linked via the child's and the socializer's interpretation of it. Consequently, studies applying the Eccles et al. model have focused on the role of previous achievement in explaining gender differences in the child's ability self-concept. These have supported the model's prediction that the child and his or her socializers both rely on the child's previous achievement-related experiences when developing gender-differentiated competence beliefs and task values (e.g., Eccles et al., 1983; Frome & Eccles, 1998; Herbert & Stipek, 2005; Tiedemann, 2000). In some empirical studies, however, the influence of prior performance on future academic outcomes was not completely mediated by motivational constructs—as proposed by the Eccles et al. model—and remained an important direct predictor of future outcomes (e.g., Watt, 2006; Zaunbauer, Retelsdorf, & Möller, 2009). The inclusion of prior achievement in future studies such as ours in which the Eccles et al. model is being applied to explain gender differences seems useful not only because of its impact on a child's ability self-concept and parents' perceptions but also because controlling for prior achievement should clarify the effects of a child's ability self-concept, task values, and parents' perceptions on gender differences in academic outcomes (cf. Eccles et al., 1983).

### **2.5 Ability Self-Concept and Task Values**

The majority of studies applying the Eccles et al. model have focused on the most proximal predictors of academic achievement and core components of the model, ability self-concept, and task values. Empirically, ability self-concept has been found to be a stronger predictor of academic achievement than of choices, and task values have been found to be more strongly associated with choices than with achievement (e.g., Berndt & Miller, 1990; Eccles et al., 1984; Meece, Wigfield, & Eccles, 1990). In congruence with common gender stereotypes, males have reported higher ability self-concept and task values in math (e.g., Eccles, Wigfield, Harold, & Blumenfeld, 1993; Huang, 2013; Steinmayr & Spinath, 2008) and females in languages and reading (e.g., Durik et al., 2006; Huang, 2013; OECD, 2002).

These gender differences in self-concepts and task values have been found to be more pronounced than the gender differences in the corresponding achievement measures (e.g., Nagy et al., 2006). However, only a few studies have tested whether gender differences in ability self-concept and task values actually *mediate* the effect of gender on academic outcomes. Eccles and colleagues (1983, 1984) found that females reported lower task values and ability self-concept in math than males, and this finding in turn predicted that females less often planned to enroll in advanced math courses. These results were recently substantiated by two longitudinal studies with large sample sizes: Nagy and colleagues (2006) and Watt (2006) found that boys' higher ability self-concept and task values in math completely explained the gender difference in the selection of advanced math courses when prior achievement was controlled for.

Results are less straightforward in the verbal domain: In two studies, ability self-concept and verbal task values partially explained girls' higher grades and better performance (Niklas & Schneider, 2012; Steinmayr & Spinath, 2008). Similarly, Eccles and colleagues (1984) found that females valued English more highly, and this in turn predicted higher numbers of women planning to study English. In other studies, however, ability self-concept and task values were not found to mediate the effect of gender on academic outcomes in the verbal domain (e.g., Durik et al., 2006; Zaunbauer et al., 2009). Taken together, the evidence shows that ability self-concept and task values mediate the effect of gender on academic outcomes in math, whereas evidence supporting such relations in the verbal domain is relatively scarce.

Considered as a whole, the Eccles et al. model is a comprehensive and complex model for explaining individual differences and gender differences in academic outcomes. Most studies that have applied the Eccles et al. model have tested the effects of a child's ability self-concept and task values and parents' perceptions of those, all of which have been found to mediate the effect of gender on academic outcomes. However, more distal factors



important for explaining individual differences in academic outcomes (e.g., child's aptitude or parents' SES) have not been considered. Furthermore, most studies, with more conclusive results, have explored gender differences in math rather than in the verbal domain. Hence, the lower achievement of boys in languages has been a relatively understudied but—against the background of boys' overall lower academic success—highly relevant research topic.

### 3. Study Overview and Hypotheses

The present study was aimed at explaining gender differences in both the mathematical and verbal domains by applying the Eccles et al. model. The current study extends prior research by including not only proximal factors such as ability self-concept and task values but also more distal factors such as the child's intelligence, one parent's perception of the child's ability, parents' education, and the child's prior achievement as potential mediators of the gender effect. We explored the relative importance of the aforementioned variables for explaining gender differences in achievement in math (in favor of boys) and German (in favor of girls). Our analyses were applied to a sample of students attending a *Gymnasium*, the highest school track in Germany, which is attended by 34% of all students and prepares them for university. On average, 54% of all *Gymnasium* students are female (Statistisches Bundesamt, 2012), indicating a greater selectivity for boys with regard to enrolling in and graduating from the *Gymnasium* than for girls (e.g., Hosenfeld, Köller, & Baumert, 1999). In previous studies using German *Gymnasium* samples, this higher selectivity for boys has often resulted in inflated mean intelligence scores for boys (e.g., Fischer, Schult, & Hell, 2013; Steinmayr & Spinath, 2008), which, first, underscores the need to include intelligence in models that are used to explain gender differences in academic achievement and, second, makes such samples especially interesting for researchers studying boys' lower grades in the verbal domain. It is possible, however, that the higher selectivity for boys might be associated not only with higher intelligence scores but also with higher SES.

To our knowledge, no one has yet tested whether gender differences in SES are related to gender differences in academic outcomes.

We expected that girls would outperform boys in German grades and that boys would outperform girls in math grades as has been found in similar samples of German Gymnasium students (e.g., Steinmayr & Spinath, 2008).

Hypothesis 1: Boys will earn better grades in math than girls.

Hypothesis 2: Girls will earn better grades in German than boys.

These gender differences in grades should be explained by gender differences in ability self-concept, task values, intelligence, prior achievement, parents' education, and parents' perception of the child's ability, indicated by a statistically significant reduction in the direct effect of gender on grades.

Hypothesis 3: The effect of gender on grades in math will be mediated by students' ability self-concept and task values in math, intelligence, prior math achievement, parents' education, and parents' perception of the child's ability in math.

Hypothesis 4: The effect of gender on grades in German will be mediated by students' ability self-concept and task values in German, intelligence, prior German achievement, parents' education, and parents' perception of the child's ability in German.

Due to the higher selectivity for boys than girls, we expected that boys would score higher on intelligence and parents' education than girls. For math, we expected that boys' higher intelligence scores and their parents' higher education as well as their higher ability self-concept, task values, prior achievement, and parents' perception of the child's math ability would contribute to and explain boys' higher grades. For the verbal domain, however, boys' higher intelligence and their parents' higher education were expected to partially compensate for girls' higher scores on ability self-concept, task values, prior achievement, and parents' perception of the child's ability in German. The total indirect effect of gender on grades via ability self-concept, task values, prior achievement, and parents' perception of the

child's ability, intelligence, and parents' education was, therefore, expected to be larger in math than in German.

Hypothesis 5: The total indirect effect of gender on grades in math will be larger than the total indirect effect of gender on grades in German.

## 4. Method

### 4.1 Sample and Procedure

The sample consisted of  $N = 520$  11th and 12th graders (age  $M = 16.94$ ,  $SD = 0.71$ ; 303 (58%) were female) from three different schools that prepare students for university (Gymnasium) located in three mid-sized towns in Germany (cf. Steinmayr, Bipp, & Spinath, 2011). Our sample can be considered representative of the typical population of this type of school (i.e., the majority being Caucasian from medium to high socioeconomic status homes). Participation was voluntary, and we received signed consent forms from parents of underage students. Almost all students were willing to participate, and only a small number of students were absent on the day of testing because they were ill, resulting in an overall participation rate of about 95%. Written questionnaires and a standardized intelligence test were administered during classes at about midterm. The intelligence test was administered using a standardized procedure by trained research assistants in groups of up to 20 students after the participants completed the questionnaire on motivation and demographics. Schools provided copies of report cards that indicated students' math and German grades, which students had received both at the end of the previous term ( $t_1$ ) and at the end of the term in which the testing took place ( $t_2$ ).

### 4.2 Materials

**Task values.** Subjective task values were assessed with three items per domain representing the three value components of the Eccles et al. model (i.e., intrinsic value, attainment value, and utility value; Steinmayr & Spinath, 2007). Students were asked to

indicate the task values they ascribed to math and German on a 5-point scale. The items were the following: “How much do you like doing math/German?”; “For me, being good in math/German is not at all/very important”; “In general, how useful is what you learn in math/German?” Cronbach’s  $\alpha$  coefficients for the combined measures were  $\alpha = .79$  for math and  $\alpha = .75$  for German. Construct validity of the measures was previously demonstrated (e.g., Steinmayr & Spinath, 2007).

**Ability self-concept.** Students’ ability self-concepts in math and German were assessed with the four items from the German “Skalen zur Erfassung des schulischen Selbstkonzepts” (SESSKO; Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, 2002). The items were “In math/German I know a little/a lot,” “To learn new things in math/German is hard/easy for me,” “In math/German, I am not talented/very talented,” and “Most assignments in math/German are hard/easy for me.” Each item was rated on a 5-point scale. Cronbach’s  $\alpha$  coefficients were  $\alpha = .95$  for math and  $\alpha = .90$  for German. The construct validity of the measures was previously demonstrated (e.g., Steinmayr & Spinath, 2007).

**Intelligence.** Participants’ intelligence was measured with the German Intelligence Structure Test 2000 R (I-S-T 2000 R; Amthauer, Brocke, Liepmann, & Beauducel, 2001). The test consists of verbal, numeric, and figural subtests. The composite score of the verbal, numeric, and figural subscale measures general reasoning and serves as an indicator of participants’ general intelligence. Internal consistencies were  $\alpha = .73$  (verbal),  $\alpha = .90$  (numeric),  $\alpha = .81$  (figural), and  $\alpha = .90$  (general reasoning).

**Academic achievement.** Students’ domain-specific academic achievement before (t1) and after testing (t2) was measured by grades in math and German. Schools provided copies of the report cards that students had received about three months before testing (t1) and about four months after testing (t2). Unfortunately, for six students at t1 and for 11 students at t2, the grades were missing either because students did not provide their names, and the report cards could not be matched to the rest of the data (t1, t2) or because students had left school in

the meantime (t2). In Germany, grades range from 1 (*excellent performance*) to 6 (*insufficient performance/fail*). To facilitate interpretation, grades were reversed so that higher values indicated better performance.

**Parents' education.** On four items, students indicated the highest scholastic and vocational degrees that their mothers and fathers had attained (cf. PISA studies; Kunter et al., 2002). The alternatives indicating parents' highest scholastic degrees ranged from 0 (*no school leaving certificate at all*) to 4 (*acquired the Abitur*; university entrance qualification awarded after successfully attending the Gymnasium). The alternatives indicating parents' highest vocational degrees ranged from 0 (*no vocational training at all*) to 5 (*acquired a PhD*).

**Parents' perception of the child's ability.** One parent of each participant was asked to fill out a questionnaire on his or her child's characteristics. A total of  $N = 376$  parents (217 mothers, 159 fathers, age  $M = 46.90$ ,  $SD = 4.96$ ) turned in the questionnaire (return rate 72%). On the questionnaire, the parent rated his or her child's abilities in math and German on a 7-point scale ranging from 1 (*very low*) to 7 (*very high*). Ability in math was defined as "solving mathematical tasks quickly and correctly and understanding new issues in math easily." Ability in German was defined as "solving tasks in German quickly and comprehensively and understanding new issues in German easily."<sup>6</sup>

### 4.3 Analyses

**Model specification.** Figure 1 shows the structural equation model (SEM) specified for testing the Eccles et al. model for explaining gender differences in academic achievement. Domain-specific task values, ability self-concept, prior achievement, intelligence, parents' education level, and parents' perception of the child's ability were included as simultaneous

---

<sup>6</sup> As an indicator of the validity of the single-item measure, participants' ability self-concept assessed with four items of SESSKO (Schöne et al., 2002) was highly correlated with participants' self-rated competencies on a single-item measure, which was identical to the item used to assess parents' perception of the child's ability ( $r = .89$  for math,  $r = .75$  for German).

potential mediators of the effect of gender on achievement. Parents' education level was indicated by fathers' and mothers' highest scholastic and vocational degree. The model allowed for covarying error terms between the fathers' and mothers' scholastic and vocational degrees because certain scholastic degrees are compulsory for certain vocational degrees (e.g., Abitur for university degrees). Participants' intelligence was indicated by the three subscales of verbal, numeric, and figural intelligence. The task values in math/German were indicated by three items representing intrinsic, attainment, and utility value. Ability self-concept in math/German was indicated by four items in each domain. Given the positive relation between domain-specific task values and ability self-concept (e.g., Eccles et al., 1983), the residuals of the two constructs were allowed to covary freely. Gender was coded as 1 = female and 2 = male, and grades at t1 and t2 and parents' perception of the child's ability were included as manifest variables. Grades at t1 were specified as predicting parents' perception of the child's ability because grades are an important source of information for parents, and parents' ratings of the child's ability were expected to rely on the child's previous grades. In line with studies that emphasized the ability of intelligence and prior achievement to predict future academic outcomes (e.g., Kuncel et al., 2004; Watt, 2006), direct paths from intelligence and grades (t1) representing students' prior achievement to grades (t2) were additionally included. Gender was specified as predicting intelligence and parents' education because of the higher selectivity for boys attending a German Gymnasium. The following cut-off scores were used to indicate a good model fit (Hu & Bentler, 1999):  $CFI \geq .95$ ,  $RMSEA \leq .06$ ,  $SRMR \leq .08$ . Standardized effects will be reported in the Results section.

Insert Figure 1 about here

**Model estimation and testing for the significance of indirect effects.** Separate SEMs for math and German were tested using the software package Mplus 7.1 (Muthén & Muthén, 1998–2013). Coefficients were estimated by applying the full information maximum

likelihood estimator (FIML), which allows cases with incomplete data to be accounted for. We calculated 95% bias-corrected bootstrap confidence intervals (95% CI) for the total, direct, and indirect effects based on 1,000 bootstrap samples to test the effects for statistical significance. Bias-corrected bootstrap confidence intervals are recommended (e.g., by Lau & Cheung, 2012, and MacKinnon, Lockwood, & Williams, 2004) because they do not assume a multivariate normal distribution of the indirect effects and are more accurate than methods that impose that assumption.

## 5. Results

Table 1 shows the means, standard deviations, and intercorrelations between all variables. For math, there were moderate to high bivariate correlations between task values, ability self-concept, parents' perception of the child's ability, and grades (range  $.52 \leq r \leq .79$ ). For German, the bivariate correlations were less pronounced, ranging from  $r = .26$  to  $r = .61$ . In order to test for overall gender differences, a MANOVA with students' gender as a factor was computed with all other variables as dependent variables. It showed a multivariate main effect of gender,  $F(13, 304) = 6.62, p < .001$ , Wilks'  $\lambda = .78$ , partial  $\eta^2 = .221$ . Univariate tests revealed the following statistically significant gender differences as depicted in Table 2: In support of Hypothesis 1, boys earned better grades in math than girls at t1 and t2. Boys also reported higher task values and ability self-concept in math than did girls.

As predicted by Hypothesis 2, girls earned better grades in German at t1 and t2. Girls also reported higher task values and ability self-concept in German than boys. Gender differences were also found in intelligence scores in favor of boys. With regard to parents' perception of the child's abilities, boys were ascribed higher competence in math and girls in German. No statistically significant gender differences were found for parents' vocational and scholastic degrees.

Insert Table 1 and 2 about here

### 5.1 Explaining Gender Differences in Math Achievement

The fit indices for the model explaining gender differences in math grades indicated a good fit: CFI = .979, RMSEA = .044, and SRMR = .040. Altogether, the model explained 55% of the variance in math grades. Hypothesis 3 predicted that the variables of the Eccles et al. model would reduce the magnitude of the direct effect of gender on math grades (with a positive algebraic sign identifying boys as earning better grades than girls). In support of Hypothesis 3, the analysis revealed a statistically significant total indirect effect of gender on grades mediated by all variables included in the model (see Table 3). However, the direct effect of gender on math grades in favor of boys was not merely reduced in magnitude but was actually reversed,  $\beta = -.10$ ,  $p = .007$ . Thus, the opposite algebraic signs of the direct and total indirect effect of gender on math grades indicate a suppression effect (e.g., MacKinnon, Krull, & Lockwood, 2000; see Table 3). When controlling for prior math achievement, parents' education, parents' perception of the child's math ability, intelligence, math ability self-concept, and task values in math, boys earned worse grades in math than girls did (see Figure 2). To gain further insights into the variables operating as suppressors, the specific indirect effects were tested for statistical significance. Three specific indirect effects were found to be statistically significant ( $\alpha = .05$ ). First, there was a specific indirect effect via boys' higher intelligence and higher prior math achievement, .06, 95% CI [.032, .087]. A second specific indirect effect was mediated by boys' higher intelligence, higher prior math achievement, and higher math values, .02, 95% CI [.002, .038]. Finally, there was a third specific indirect effect via boys' higher intelligence, higher prior math achievement, parents' higher math ability perception of boys, and boys' higher math values, .016, 95% CI [.001, .03]. Consequently, task values in math, parents' perception of child's math ability, previous math achievement, and intelligence were found to operate as suppressors in this model, reversing the direction of the gender difference in math grades. All other specific indirect effects did not differ significantly from zero.



Insert Table 3 and Figure 2 about here

## 5.2 Explaining Gender Differences in German Achievement

The fit indices of the model explaining gender differences in German grades indicated an acceptable fit: CFI = .949, RMSEA = .056, and SRMR = .049. Overall, the model explained 30% of the variance in German grades. Hypothesis 4 predicted that the variables of the Eccles et al. model would also reduce the magnitude of the direct effect of gender on grades in German (with girls earning higher grades than boys). Even when controlling for prior German achievement, parents' education, parents' perception of child's German ability, intelligence, German ability self-concept, and task values in German, the direct effect of gender on German grades was not statistically significantly reduced as indicated by the overlapping confidence intervals of the total and direct effects (see Table 3 and Figure 3). The direct effect remained statistically significant. Further, the total indirect effect via all mediators in the model was not statistically significant, so Hypothesis 4 was not supported by the data. Closer inspection of the specific indirect effects revealed two opposing effects: On the one hand, girls' higher grades could be partially explained by their better prior achievement, with a specific indirect effect of  $-.074$ , 95% CI  $[-.11, -.037]$ . On the other hand, this effect was canceled out by a suppression effect, a specific indirect effect with a positive sign indicating higher performance for boys than girls: Boys' higher intelligence positively predicted prior achievement, which in turn predicted the higher achievement of boys at  $t_2$ ,  $.021$ , 95% CI  $[.001, .04]$ . The two specific indirect effects seemed to neutralize each other, thus resulting in a nearly unchanged direct effect of gender on grades. All other specific indirect effects were not significantly different from zero.

Insert Figure 3 about here

## 5.3 Comparison of the Results for Math and German

As depicted in Table 3, the total indirect effect of gender on math grades was found to be larger than the one of gender on German grades. Because the confidence intervals of the

indirect effects in the two domains did not overlap, the difference in magnitude was found to be statistically significant, and Hypothesis 5 was supported by our data.

Taken together, prior math achievement, parents' education, parents' perception of child's math ability, intelligence, math ability self-concept, and task values in math suppressed the actual association between gender and math grades. When these variables were controlled, the sign of the direct effect of gender changed from positive (indicating higher grades for boys) to negative (indicating higher grades for girls). This means that when gender differences in these variables were taken into account, boys also earned lower grades in math than girls. More precisely, we found three statistically significant suppression effects via intelligence, prior math achievement, task values in math, and parents' perception of child's math ability. Contrary to the results for math, prior German achievement, parents' education, parents' perception of child's German ability, intelligence, German ability self-concept, and valuing German did not explain girls' higher achievement in German. Closer inspection revealed two specific indirect effects with opposing signs for German so that the mediation effect via girls' higher previous achievement was canceled out by the suppression effect from boys' higher intelligence predicting the higher achievement of boys than girls at t1 and t2.

## 6. Discussion

The present study aimed to explain gender differences in both the mathematical and the verbal domains by applying the Eccles et al. model. In line with Hypotheses 1 and 2, boys outperformed girls in math grades and vice versa in German grades. In order to explain these gender differences using structural equation modeling, we simultaneously tested prior achievement, parents' education, parents' perception of child's ability, intelligence, ability self-concept, and task values as potential mediators of the effect of gender on grades in math and German. Different results were found in the two domains. As predicted by Hypothesis 3, the mentioned variables mediated the effect of gender on math grades. However, the effect of

gender on math grades was not just reduced in magnitude but was actually reversed: When important motivational, cognitive, and socialization influences were controlled for, boys were found to earn lower grades than girls in math, indicating a suppression effect. Closer inspection identified that intelligence, prior achievement in math, task values in math, and parents' perception of child's math ability operated as suppressors in the math model. Girls' higher grades in German, however, were not explained by the included variables, and no statistically significant total indirect effect was found (i.e., Hypothesis 4 was rejected). Moreover, the analyses showed that the specific indirect effect that was mediated via girls' higher previous achievement in German was canceled out by the specific indirect effect via boys' higher intelligence, which predicted greater previous achievement. In line with Hypothesis 5, the indirect effect of gender on grades was found to be larger in math than in German.

Taken together, the first important result of our study was that boys' higher grades in math were not simply explained but were rather over-explained by the variables from the Eccles et al. model that we included in our analyses. When prior achievement, parents' education, parents' perception of child's ability, intelligence, ability self-concept, and task values were controlled for, boys were found to earn *lower* grades than girls—even in math. Although boys had a greater ability self-concept in math than girls, ability self-concept did not mediate the effect of gender on math grades. However, the lack of mediation via ability self-concept should not be strongly interpreted because math ability self-concept was substantially correlated with math task values, parents' perception of child's math ability, and prior math achievement, which, along with intelligence, mediated the effect of gender on performance (see Table 1; for a discussion on complementary effects and competition among mediators, cf. Hayes, 2013).

Our finding that girls also outperform boys in math when important motivational, cognitive, and socialization influences are controlled for casts a different light on recent meta-

analytic results that reported no gender differences in math grades for non-North-American samples (Voyer & Voyer, 2014) by unmasking an underperformance of boys even in this still male-stereotyped domain (e.g., Steffens et al., 2010). Boys' lower success in turning their potential into high grades—as reflected in our results for language *and* math—is likely to contribute to their currently overall lower academic participation and success, such as their lower GPAs (NCES, 2004) and fewer higher level school leaving certificates (e.g., in Germany: Statistisches Bundesamt, 2012). Further, it matches the results of previous studies that have reported that boys tend to earn lower grades than what would be expected from their scores on competence or aptitude tests, whereas girls' test scores tend to underpredict their grades (e.g., Duckworth & Seligman, 2006; Neumann, Milek, Maaz, & Gresch, 2010; Valtin, Wagner, Schwippert, 2005). For example, in a sample comparable to the present study's sample, adolescent girls performed better in math (indicated by grades) than boys when intelligence alone was controlled for (Steinmayr & Spinath, 2008).

The second important finding of our study is that the variables included in our analyses, namely prior achievement, parents' education, parents' perception of child's ability, intelligence, ability self-concept, and task values, could *not* explain why boys earned lower grades than girls in the verbal domain. As in the math model, gender was found to be a direct predictor of grades in the language model as well, a finding that reveals an underachievement of boys across both domains that cannot be explained by the mentioned variables when they are included as mediators. This suggests that boys' underachievement may be traced back to *domain-unspecific* variables and mechanisms that lie outside the part of the Eccles et al. model that we applied (i.e., domain-specific constructs such as task values in math or language). Drawing on results from other strands of research, variables such as boys' lower effort and engagement and less adjusted behavior in class appear to be crucial for explaining boys' lower academic success. Steinmayr and Spinath (2008), for instance, found that girls' higher agreeableness and lower work avoidance partially explained girls' higher GPAs.

Similar results have been obtained for girls' higher social and behavioral skills (DiPrete & Jennings, 2012), better classroom behavior, greater effort (Downey & Vogt Yuan, 2005), better self-discipline (Duckworth & Seligman, 2006), and stronger commitment and willingness to learn (Neugebauer, 2011).

For explaining why boys in general show lower effort and less adjusted behavior in school than girls, the construct of "values" as it is conceptualized in expectancy-value theories (e.g., Eccles, 2007) seems of crucial importance. Gender differences in academic engagement might be due to gender differences in the attainment value and costs attached to academic engagement and school in general. Expectancy-value theory predicts that the attainment value of an activity is high if engaging in it is consistent with one's identity (Eccles, 2007). In previous studies, the masculine image of STEM subjects, which clashes with a girl's femininity, was identified as one factor that could explain why girls stay away from STEM subjects (e.g., Steffens et al., 2010; for an overview, see Kessels, Heyder, Latsch, & Hannover, 2014). Because school in general and showing effort and diligence in the school context are perceived as feminine (Heyder & Kessels, 2013; Jackson, 2002; 2003), they are likely to be perceived as more consistent with a girl's than with a boy's identity. This lower fit in turn is assumed to explain why boys value school in general and effort in the school context less than girls (Kessels et al., 2014; cf. OECD, 2013b). Further, showing effort in school seems to have more costs for boys than for girls: Boys who work hard in school are at risk of being perceived as less masculine or even as feminine, and such perceptions have been proposed to make them less popular among their peers (Jackson, 2002; 2003). Girls, on the contrary, seem to benefit from a better fit between the culturally shared conception of femininity (as compared with masculinity) and academic demands (e.g., Burke, 1989; Kessels & Steinmayr, 2013; Mickelson, 1989) as this better fit is supposed to result in the greater academic engagement of girls than boys (Kessels et al., 2014). Taken together, our results show that boys realize their potential across domains in academic success to a lesser extent

than girls (cf. Freudenthaler, Spinath, & Neubauer, 2008). Second, our results indicate the importance of including the values and costs attached to school and academic engagement in general in future studies that are geared toward explaining gender differences in academic achievement.

Limitations of our study concern the correlational and mostly cross-sectional nature of our data, which restricts causal inferences. Only students' grades were assessed longitudinally, at about three months before and four months after the testing took place. Therefore, we cannot exclude the possibility that the variables considered in our analyses as mediators might have already led to gender differences in prior achievement in the first place. The correlations between the variables included as mediators further limit insights into which portion of the effect could be traced back to which particular construct (cf. Hayes, 2013). Furthermore, our study was limited to a sample of German adolescents attending the *Gymnasium*, the highest school track in Germany. In our sample, boys scored higher than girls on the intelligence test, caused by the greater selectivity for boys enrolled in a *Gymnasium*, replicating the findings of other German studies (cf. Fischer et al., 2013; Steinmayr & Spinath, 2008). Given the predictive power of intelligence for academic achievement (e.g., Deary et al., 2007; Kuncel et al., 2004), different relations might be observed in less selective samples. In previous studies, the higher selectivity for boys in *Gymnasiums* has been used to explain why gender differences in math competence in favor of boys have been found within each different school track in Germany, but such effects diminish if the whole sample is analyzed (Hosenfeld et al., 1999). Similarly, the higher selectivity for boys may also explain why in our study and in some other German studies (e.g., Dresel et al., 2006; Dumont et al., 2012), boys have been found to earn better grades in math than girls at first glance, whereas in a recent meta-analysis, no gender differences in math grades in studies from countries outside of North America were found, and small

advantages in favor of girls were found in North-American studies (Voyer & Voyer, 2014). Beyond that, different results may be obtained for other age groups.

Previous studies that have aimed to explain gender differences in academic outcomes have applied smaller parts of the Eccles et al. model, most commonly within the domain of math. Our study extended prior research by including not only ability self-concepts and values but also the child's intelligence, parents' perception of the child's ability, socioeconomic status, and the child's prior achievement as potential mediators of the gender effect on math and language achievement. As our study showed that gender differences in math achievement resemble those in language achievement when important predictors of academic success are controlled for, future research on gender differences in academic outcomes may gain new insights from our study when focusing on domain-unspecific motivational constructs and behavior such as engagement and social skills. Within the framework of the Eccles et al. model, it seems fruitful to test how values and costs attached to school in general can foster or hinder the display of academic engagement. In addition, longitudinal studies comprising motivational, cognitive, and sociocultural factors are needed to clarify the causal order and to disentangle the effects of the individual factors predicted by the Eccles et al. model.

### References

- Amthauer, R., Brocke, B. L. D., & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R* [Intelligence-Structure-Test 2000 R]. Göttingen: Hogrefe.
- Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
- Berndt, T. J., & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology, 82*, 319–326. doi:10.1037/0022-0663.82.2.319
- Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*, 319–337. doi:10.1146/annurev.soc.34.040507.134719
- Burke, P. J. (1989). Gender identity, sex, and school performance. *Social Psychology Quarterly, 52*, 159–169. Retrieved from <http://www.jstor.org/stable/2786915>
- Calvin, C. M., Fernandes, C., Smith, P., Visscher, P. M., & Deary, I. J. (2010). Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. *Intelligence, 38*, 424–432. doi:10.1016/j.intell.2010.04.005
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13–21. doi:10.1016/j.intell.2006.02.001
- DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research, 41*, 1–15. doi:10.1016/j.ssresearch.2011.09.001
- Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly, 46*, 299–321. doi:10.1111/j.1533-8525.2005.00014.x
- Dresel, M., Stöger, H., & Ziegler, A. (2006). Klassen- und Schulunterschiede im Ausmaß von Geschlechterdiskrepanzen bei Leistungsbewertungen und Leistungsaspirationen: Ergebnisse einer Mehrebenenanalyse [Class and school differences in gender discrepancies in performance evaluations and achievement aspirations: Results of a multilevel analysis]. *Psychologie in Erziehung und Unterricht, 53*, 44–61.
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198–208. doi:10.1037/0022-0663.98.1.198
- Dumont, H., Trautwein, U., Lüdtke, O., Neumann, M., Niggli, A., & Schnyder, I. (2012). Does parental homework involvement mediate the relationship between family background and educational outcomes? *Contemporary Educational Psychology, 37*, 55–69. doi:10.1016/j.cedpsych.2011.09.004
- Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*, 382–393. doi:10.1037/0022-0663.98.2.382
- Eccles, J. S. (2007). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). New York: Guilford Press.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *A Series of books in psychology. Achievement and achievement motives. Psychological and sociological approaches* (pp. 76–146). San Francisco: W.H. Freeman.



- Eccles, J. S., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, *46*, 26–43. doi:10.1037/0022-3514.46.1.26
- Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*, 183–201. doi:10.1111/j.1540-4560.1990.tb01929.x
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. doi:10.1146/annurev.psych.53.100901.135153
- Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*, 830–847. doi:10.2307/1131221
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education*, *80*, 114–138. doi:10.1177/003804070708000202
- Fischer, F., Schult, J., & Hell, B. (2013). Sex differences in secondary school success: why female students perform better. *European Journal of Psychology of Education*, *28*, 529–543. doi:10.1007/s10212-012-0127-4
- Freudenthaler, H. H., Spinath, B., & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality*, *22*, 231–245. doi:10.1002/per.678
- Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, *74*, 435–452. doi:10.1037/0022-3514.74.2.435
- Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, *33*, 9–33. doi:10.1016/j.cedpsych.2006.12.002
- Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York: Psychology Press.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: The Guilford Press.
- Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, *26*, 276–295. doi:10.1016/j.appdev.2005.02.007
- Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69*, 605–617. doi:10.1007/s11199-013-0309-9
- Hosenfeld, I., Köller, O., & Baumert, J. (1999). Why sex differences in mathematics achievement disappear in German secondary schools: A reanalysis of the German TIMSS-data. *Studies in Educational Evaluation*, *25*, 143–161. doi:10.1016/S0191-491X(99)00018-8
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. doi:10.1080/10705519909540118
- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*, 505–528. doi:10.1016/j.jsp.2011.07.001

- Huang, C. (2013). Gender differences in academic self-efficacy: a meta-analysis. *European Journal of Psychology of Education*, 28, 1–35. doi:10.1007/s10212-011-0097-y
- Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy. *Gender and Education*, 14, 37–50. doi:10.1080/09540250120098870
- Jackson, C. (2003). Motives for 'laddishness' at school: Fear of failure and fear of the 'feminine'. *British Educational Research Journal*, 29, 583–598. doi:10.1080/01411920301847
- Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, 83, 518–527. doi:10.1037/0022-0663.83.4.518
- Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J., & Malanchuk, O. (2005). I can, but I don't want to: The impact of parents, interests, and activities on gender differences in math. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics. An integrative psychological approach* (pp. 246–263). Cambridge, UK, New York: Cambridge University Press.
- Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M., & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, 42, 11–26. doi:10.1037/0012-1649.42.1.11
- Kessels, U., Heyder, A., Latsch, M., & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, 56, 219–228. doi:10.1080/00131881.2014.898916
- 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. doi:10.1016/j.lindif.2012.09.013
- Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, 105, 198–214. doi:10.1037/0033-2909.105.2.198
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148–161. doi:10.1037/0022-3514.86.1.148
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., ... Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente* [PISA 2000: Documentation of assessment instruments]. Berlin: Max-Planck-Inst. für Bildungsforschung.
- Kyttälä, M., & Björn, P. M. (2010). Prior mathematics achievement, cognitive appraisals and anxiety as predictors of Finnish students' later mathematics performance and career orientation. *Educational Psychology*, 30, 431–448. doi:10.1080/01443411003724491
- Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods*, 15, 3–16. doi:10.1177/1094428110391673
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, 1, 173–181. doi:10.1023/A:1026595011371
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99–128. doi:10.1207/s15327906mbr3901\_4

- Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*, 60–70. doi:10.1037/0022-0663.82.1.60
- Mensah, F. K., & Kiernan, K. E. (2010). Gender differences in educational attainment: influences of the family environment. *British Educational Research Journal, 36*, 239–260. doi:10.1080/01411920902802198
- Mickelson, R. A. (1989). Why does Jane read and write so well? The anomaly of women's achievement. *Sociology of Education, 62*, 47–63.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. Boston: TIMSS & PIRLS International Study Center.
- Muthén, L. K. & Muthén, B. O. (1998-2013). Mplus Version 7.1. Los Angeles, CA: Muthén & Muthén.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12*, 323–345. doi:10.1080/13803610600765687
- National Center for Education Statistics. (2004). *The high school transcript study: A decade of change in curricula and achievement, 1990-2000*. Washington, DC: U.S. Government Printing Office.
- National Center for Education Statistics. (2009). *The condition of education*. Washington, DC: U.S. Government Printing Office.
- Neisser, U., Boodoo, G., Bouchard, T. J., JR., Boykin, A. W., Brody, N., Ceci, S. J., ... Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77–101. doi:10.1037/0003-066X.51.2.77
- Neugebauer, M. (2011). Werden Jungen von Lehrerinnen bei den Übergangsempfehlungen für das Gymnasium benachteiligt? Eine Analyse auf Basis der IGLU-Daten [Do female teachers discriminate against boys in tracking recommendations? An analysis on the basis of the IGLU-data]. In A. Hadjar (Ed.), *Werden Jungen von Lehrerinnen bei den Übergangsempfehlungen für das Gymnasium benachteiligt? Eine Analyse auf Basis der IGLU-Daten* (pp. 235–260). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Neumann, M., Milek, A., Maaz, K., & Gresch, C. (2010). Zum Einfluss der Klassenzusammensetzung auf den Übergang von der Grundschule in die weiterführenden Schulen [On the influence of class composition on the transition from primary to secondary schools]. In Bundesministerium für Bildung und Forschung (Ed.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (pp. 229–252). Berlin: Bundesministerium für Bildung und Forschung.
- Niklas, F., & Schneider, W. (2012). Die Anfänge geschlechtsspezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen [The beginning of gender-based performance differences in mathematics and linguistic competencies]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 44*, 123–138. doi:10.1026/0049-8637/a000064
- OECD. (2002). *Reading for Change: Performance and Engagement across Countries: Results from PISA 2000*. Paris: OECD Publishing.
- OECD. (2013a). *PISA 2012 Results: What students know and can do (Volume I)*. Paris: OECD Publishing.

- OECD. (2013b). *PISA 2012 Results: Ready to learn (Volume III)*. Paris: OECD Publishing.
- Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology, 38*, 225–235. doi:10.1016/j.cedpsych.2013.03.004
- Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). *SESSKO Skalen zur Erfassung des schulischen Selbstkonzept* [SESSKO Scales for the Assessment of School-Related Competence Beliefs]. Göttingen: Hogrefe.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417–453. doi:10.3102/00346543075003417
- Spinath, B., Freudenthaler, H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences, 48*, 481–486. doi:10.1016/j.paid.2009.11.028
- Statistisches Bundesamt. (2012). *Bildung und Kultur: Allgemeinbildende Schulen Schuljahr 2011/2012* [Education and culture: General schools schoolyear 2011/2012]. Wiesbaden: Statistisches Bundesamt.
- Steffens, M. C., 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. doi:10.1037/a0019920
- Steinmayr, R., Bipp, T., & Spinath, B. (2011). Goal orientations predict academic performance beyond intelligence and personality. *Learning and Individual Differences, 21*, 196–200. doi:10.1016/j.lindif.2010.11.026
- Steinmayr, R., & Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift für Pädagogische Psychologie, 21*, 207–216. doi:10.1024/1010-0652.21.3.207
- 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. doi:10.1002/per.676
- Steinmayr, R., Dinger, F.C., & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality, 26*, 335–349. doi: 10.1002/per.842
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology, 92*, 144–151. doi:10.1037/0022-0663.92.1.144
- Valtin, R., Wagner, C., & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse- schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen [Students at the end of fourth grade- performance, attitudes, and outer school learning conditions]. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin, & G. Walther (Eds.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (pp. 187–238). Münster [u.a.]: Waxmann.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*, 1174–1204. doi:10.1037/a0036620

- Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educational Research and Evaluation, 12*, 305–322. doi:10.1080/13803610600765562
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461–481. doi:10.1037/0033-2909.91.3.461
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. doi:10.1006/ceps.1999.1015
- Zaubauer, A. C. M., Retelsdorf, J., & Möller, J. (2009). Die Vorhersage von Englischleistungen am Anfang der Sekundarstufe [Prediction of English achievement in early secondary school]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41*, 153–164. doi:10.1026/0049-8637.41.3.153

Table 1  
Means, standard deviations, and bivariate correlations of all variables

	<i>M</i>	<i>SD</i>	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Gender			<b>.22</b>	<b>-.29</b>	<b>.29</b>	<b>-.28</b>	<b>.31</b>	<b>.12</b>	<b>-.15</b>	<b>.10</b>	<b>-.13</b>	<b>.07</b>	<b>.06</b>	<b>.24</b>	<b>-.17</b>
2 Task values in math	3.30	0.97		<b>-.12</b>	<b>.79</b>	<b>-.30</b>	<b>.25</b>	<b>.53</b>	<b>.04</b>	<b>.57</b>	<b>.08</b>	<b>.10</b>	<b>.08</b>	<b>.60</b>	<b>-.12</b>
3 Task values in German	3.70	0.79			<b>-.28</b>	<b>.61</b>	<b>-.18</b>	<b>-.14</b>	<b>.27</b>	<b>-.20</b>	<b>.19</b>	<b>-.01</b>	<b>-.02</b>	<b>-.20</b>	<b>.32</b>
4 Math ability self-concept	3.15	1.03				<b>-.36</b>	<b>.41</b>	<b>.64</b>	<b>.01</b>	<b>.64</b>	<b>.10</b>	<b>.13</b>	<b>.13</b>	<b>.73</b>	<b>-.19</b>
5 German ability self-concept	3.66	0.78					<b>-.14</b>	<b>-.25</b>	<b>.35</b>	<b>-.27</b>	<b>.26</b>	<b>-.00</b>	<b>.00</b>	<b>-.26</b>	<b>.46</b>
6 Intelligence	109.46	17.79						<b>.33</b>	<b>.11</b>	<b>.34</b>	<b>.15</b>	<b>.23</b>	<b>.18</b>	<b>.42</b>	<b>.03</b>
7 Math grade (t1)	3.93	1.05							<b>.30</b>	<b>.64</b>	<b>.20</b>	<b>.18</b>	<b>.15</b>	<b>.60</b>	<b>-.00</b>
8 German grade (t1)	4.17	0.80								<b>.18</b>	<b>.45</b>	<b>.12</b>	<b>.14</b>	<b>.16</b>	<b>.43</b>
9 Math grade (t2)	3.88	1.14									<b>.29</b>	<b>.18</b>	<b>.21</b>	<b>.52</b>	<b>-.05</b>
10 German grade (t2)	4.08	0.78										<b>.19</b>	<b>.29</b>	<b>.10</b>	<b>.27</b>
11 Parents' scholastic degree	2.44	0.94											<b>.75</b>	<b>.13</b>	<b>.03</b>
12 Parents' vocational degree	2.06	1.02												<b>.15</b>	<b>.08</b>
13 Parents' perception of child's math ability	4.72	1.40													<b>.07</b>
14 Parents' perception of child's German ability	5.22	1.07													

Note.  $N = 376-520$ . Gender was coded 1 = female, 2 = male. Task values and ability self-concept range from 1 to 5 with 5 indicating higher values in the direction of the scale. Intelligence scores are standard values ( $M = 100$ ,  $SD = 10$ ). Grades range from 1 to 6 with higher values indicating better performance. Parents' scholastic degrees range from 0 (no school leaving certificate at all) to 4 (acquired the *Abitur*; university entrance qualification). Parents' vocational degrees range from 0 (no vocational training at all) to 5 (acquired a *PhD*). Parents' perception of the child's abilities range from 1 to 7 with higher values indicating higher ability perceptions. Statistically significant correlations are bold with,  $r \geq |.10|$ ,  $p \leq .05$ ;  $r \geq |.12|$ ,  $p \leq .01$ ;  $r \geq |.17|$ ,  $p \leq .001$ .

Table 2  
Results of univariate tests for gender differences

	Girls			Boys			<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>				
Task values in math	3.12	0.96	303	3.55	0.93	217	-5.04	518	.000	0.45
Task values in German	3.89	0.70	303	3.43	0.83	217	6.60	414.96	.000	0.60
Math ability self-concept	2.89	0.98	303	3.50	0.99	217	-6.92	518	.000	0.62
German ability self-concept	3.84	0.72	303	3.41	0.78	217	6.50	518	.000	0.58
Intelligence	104.80	17.11	303	115.97	16.69	217	-7.42	518	.000	0.66
Math grade (t1)	3.82	1.01	301	4.08	1.10	213	-2.73	512	.007	0.24
German grade (t1)	4.27	0.76	302	4.03	0.83	213	3.50	513	.001	0.31
Math grade (t2)	3.79	1.11	297	4.02	1.18	212	-2.31	507	.022	0.21
German grade (t2)	4.16	0.77	297	3.95	0.78	212	3.01	507	.003	0.27
Parents' scholastic degree	2.39	0.93	301	2.52	0.96	216	-1.59	515	.112	0.14
Parents' vocational degree	2.01	0.98	251	2.13	1.07	185	-1.17	337.08	.244	0.12
Parents' perception of child's math ability	4.47	1.36	238	5.17	1.37	138	-4.82	374	.000	0.52
Parents' perception of child's German ability	5.35	1.01	237	4.98	1.13	137	3.31	372	.001	0.36

*Note.* Task values and ability self-concept range from 1 to 5 with 5 indicating higher values in the direction of the scale. Intelligence scores are standard values ( $M = 100$ ,  $SD = 10$ ). Grades range from 1 to 6 with higher values indicating better performance. Parents' scholastic degrees range from 0 (*no school leaving certificate at all*) to 4 (*acquired the Abitur*; university entrance qualification). Parents' vocational degrees range from 0 (*no vocational training at all*) to 5 (*acquired a PhD*). Parents' perception of the child's abilities range from 1 to 7 with higher values indicating higher ability perceptions.

Table 3

*Standardized estimates of the effects of gender on grades in math and German (t2)*

	Total effect	Direct effect	Total indirect effect
	[95% CI]	[95% CI]	[95% CI]
Math	.10	-.10	.20
	[.016; .187]	[-.177; -.028]	[.127; .282]
German	-.14	-0.12	-0.02
	[-.223; -.049]	[-.213; -.017]	[-.100; .059]

*Note.* Positive signs indicate better performance of boys. Negative signs indicate better performance of girls. 95% CI = 95% bias-corrected bootstrap confidence interval.



Figure 1. The structural equation model specified for testing the Eccles et al. model for explaining gender differences in academic achievement. Parents' education is indicated by fathers' and mothers' scholastic and vocational degrees (Sch\_f, V\_f, Sch\_m, V\_m). Intelligence is indicated by three subscales (numerical, figural, verbal). Task values are comprised of intrinsic value (IV), attainment value (AV), and utility value (UV). Ability self-concept is indicated by four items (ASC1, ASC2, ASC3, ASC4).

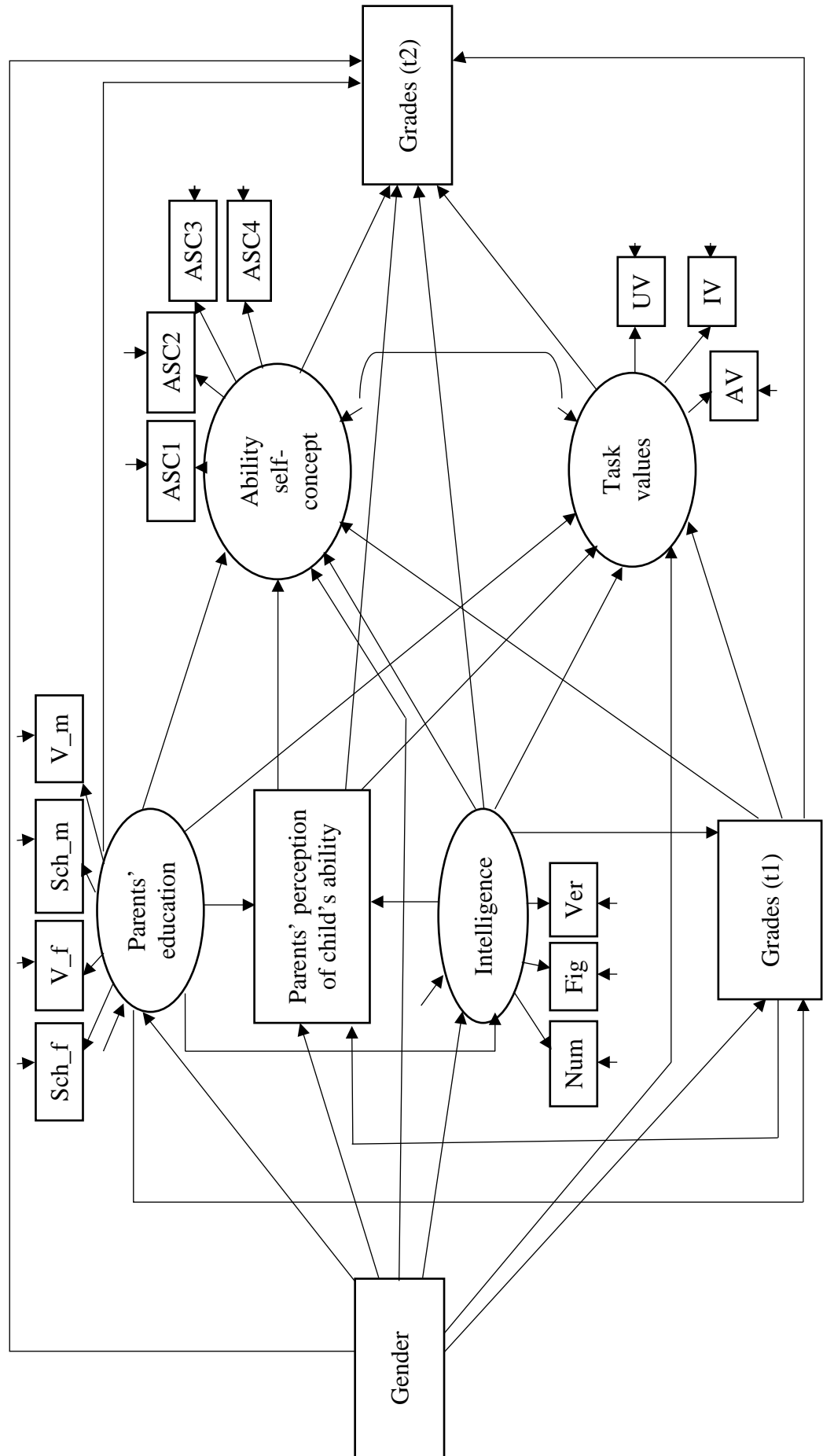


Figure 2. Results of the SEM specified for testing the Eccles et al. model for explaining gender differences in math grades. Standardized coefficients for specified paths are presented. Dotted arrows represent nonsignificant paths ( $p \geq .05$ ). Gender was coded 1 = female, 2 = male.

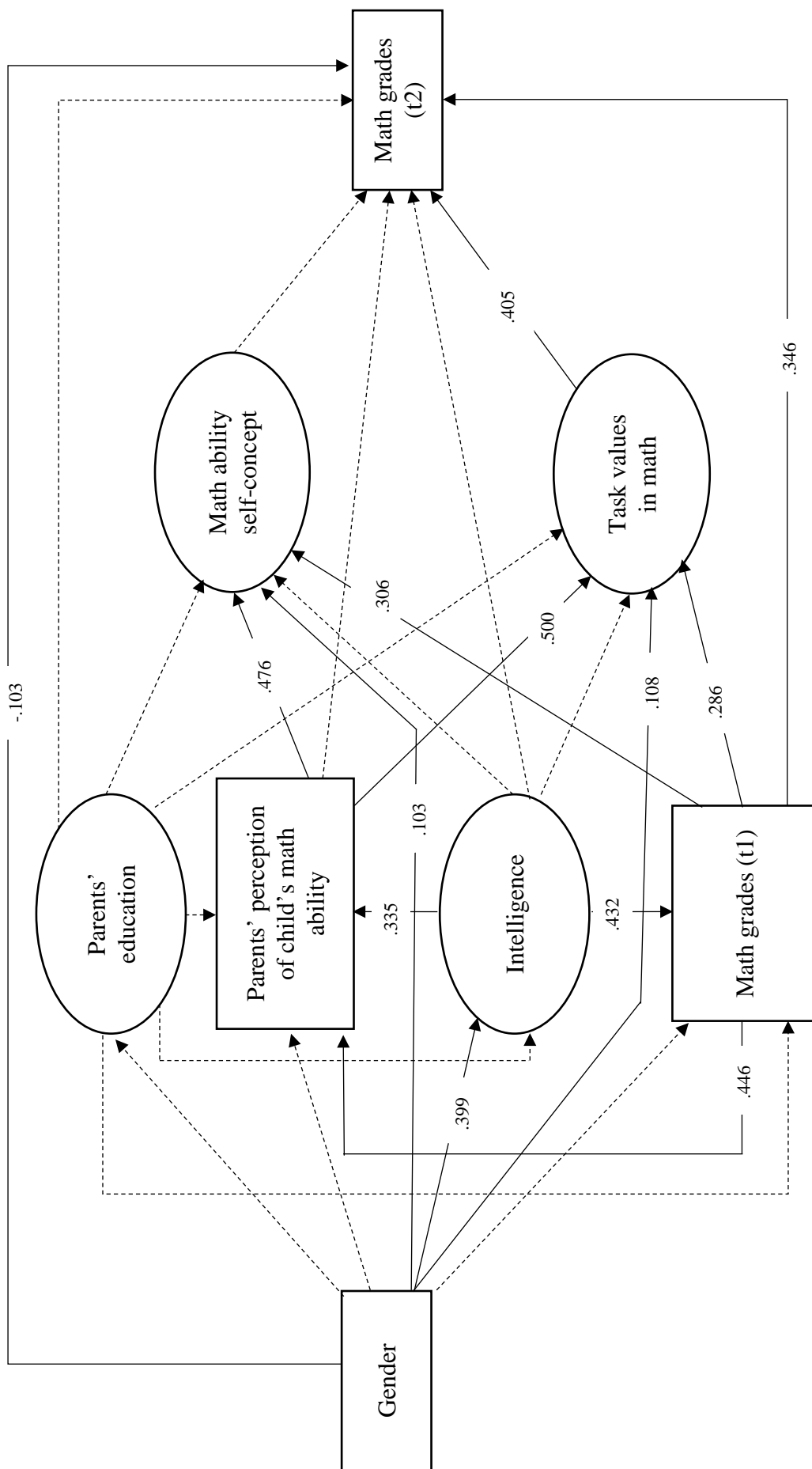
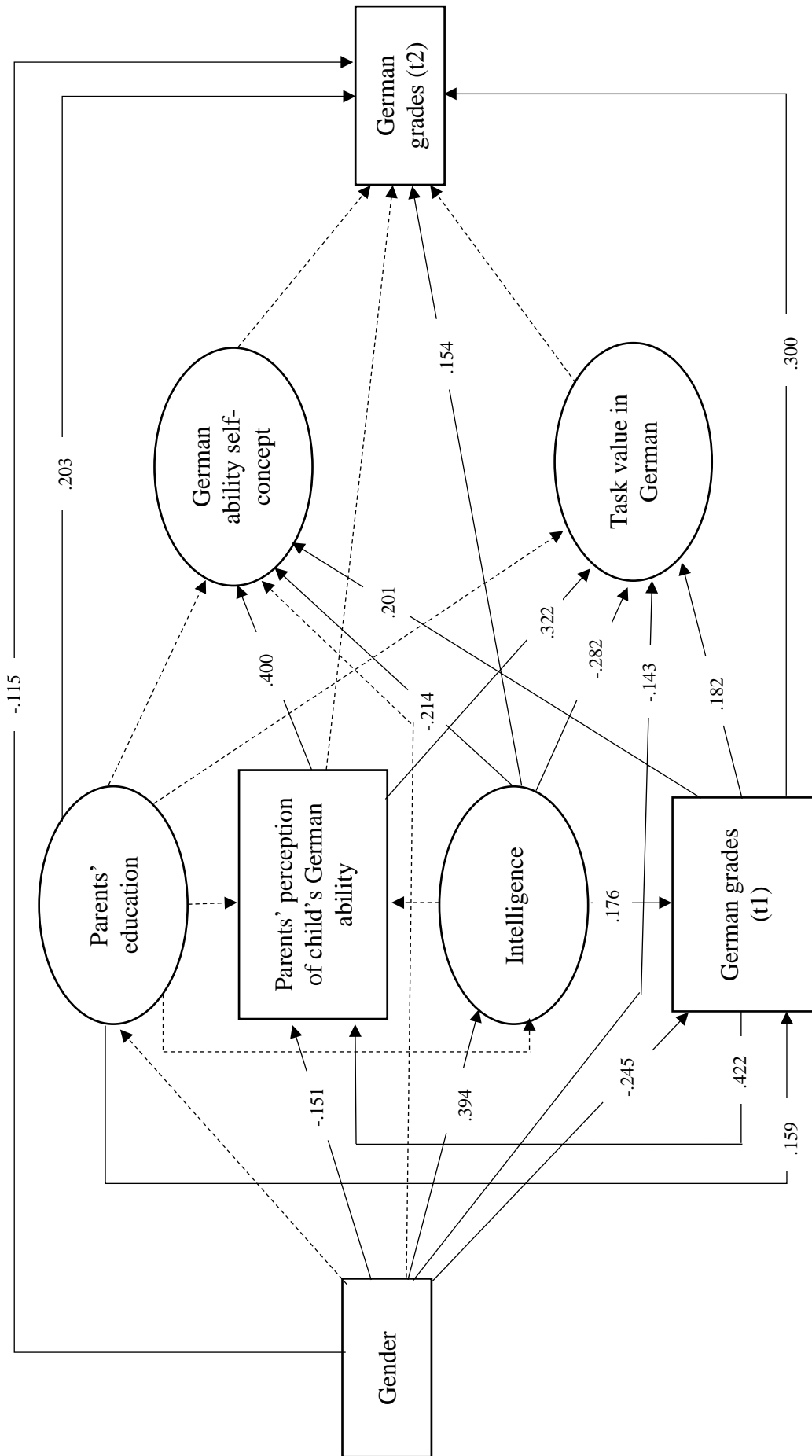


Figure 3. Results of the SEM specified for testing the Eccles et al. model for explaining gender differences in German grades. Standardized coefficients for specified paths are presented. Dotted arrows represent nonsignificant paths ( $p \geq .05$ ). Gender was coded 1 = female, 2 = male.





## **5.2 Studie 2: Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement**

Note: This is the first author's version of a work that was published in *Sex Roles*. The final publication is available at Springer via <http://dx.doi.org/10.1007/s11199-013-0309-9>.

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69*, 605–617. doi:10.1007/s11199-013-0309-9

The current research was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) allocated to the second author (KE 1412/2-1).

### Abstract

One cause proposed for boys' relatively lower academic achievement is a "feminisation" of schools that might result in a lack of fit between boys' self-concept and academic engagement. Research so far has investigated math-male and language-female stereotypes, but no school-female stereotypes. Our study tested for implicit gender stereotyping of school and its impact on boys' achievement in  $N = 122$  ninth-graders from a large city in Western Germany using the Go/No-go Association Task (GNAT). Gender role self-concept and grades in math (representing an academic domain stereotyped as male) and German (domain stereotyped as female) were assessed using written questionnaires. It was found that, overall, students associated school more strongly with female than with male, and that this association of school with female was related to boys' academic achievement. The more strongly boys associated school with female and the more they ascribed negative masculine traits to themselves, the lower their grades in German were. Boys' academic achievement in math was unrelated to the extent to which they perceived school as feminine and themselves as masculine. Girls' grades in both German and math were unrelated to their gender stereotyping of school. These findings emphasize the importance of fit between a student's gender, gender role self-concept and gender stereotyping of school for academic achievement. Strategies to improve this fit are discussed.

Key words: gender gap in academic achievement, gender role self-concept, implicit measures, stereotypes, school

### Introduction

In many countries, male students lag behind female students on important indicators of academic success (for a comprehensive overview see Buchmann, DiPrete, & McDaniel, 2008). In the U.S., for instance, boys earn lower grades during middle and high school (e.g. Clark, Lee, Goodman, & Yacco, 2008) and graduate with lower GPAs than their female classmates (NCES, 2004). Similarly in Great Britain, academic performance in girls is found to be superior to that of boys, especially with regard to GCSE examinations (e.g. Gray, Peng, Steward, & Thomas, 2004), for which the achievement gap widened from 1992/1993 through 2002/03 (Lindsay & Muijs, 2006). In Germany today, both male and female students earn a larger number of higher-level school leaving certificates as compared with earlier generations, but in this, as well, female students are pulling ahead of their male classmates, with 46% of young women and 37% of young men obtaining the *Abitur*, the qualification for entry to university (Hannover & Kessels, 2011). In line with this, girls also earn better grades than boys in Germany, especially for the subject German (for primary school e.g. Helbig, 2010; for secondary school e.g. Steffens, Jelenec, & Noack, 2010). For math grades, the picture is less consistent with most studies reporting no gender differences (for primary school e.g. Helbig, 2010; for secondary school e.g. Köller, Daniels, Schnabel, & Baumert, 2000; Steffens et al., 2010) and a few reporting relatively small advantages for boys in primary school (e.g. Tiedemann & Faber, 1994).

The gender gap in academic success is not fully reflected in gender differences found on standardized tests. While girls outperform boys in reading literacy in many countries and boys demonstrate greater competency in math and/or science (OECD, 2009), studies including both grades and standardized performance measures have shown that boys on average earn lower grades than expected from their performance on ability or achievement tests (Duckworth & Seligman, 2006), so that the current gender gap in academic success

seems not due to any lower cognitive potential in boys but rather to non-cognitive variables that are also important for earning high grades.

Taken together, many studies from different Western, industrialized countries currently show relatively lower academic success for boys as compared with girls. Using a German sample of adolescents, the present study seeks to add to our understanding of why this is the case. In brief, we test whether boys' lower grades can be traced back to a misfit between their perception of school as something feminine and their self-perception as masculine. In line with this, we bring together several strands of research: first, studies of (explicit and implicit) gender stereotyping of academic domains and its impact on students' motivation and performance; second, studies that discuss several aspects of a possible, general "feminisation" of school; and, third, studies that show boys themselves might regard academic engagement as unsuited to their masculinity. Many of these last studies were based on qualitative interviews. In our study we aim at combining this rich body of research with measures used in social psychology studies of gender stereotypes. Specifically, we use an implicit measure (Go/No-go Association Task, Nosek & Banaji, 2001) that assesses the automatic, non-conscious gender stereotyping of a given construct, in our case, school.

### **Gender Stereotyping of School Subjects**

In this section, we review evidence as to what extent specific school subjects or domains are perceived as gender stereotyped. We report both evidence using "explicit measures" such as self-report questionnaires and findings based on "implicit measures" using reaction time latencies in order to assess automatic, implicit gender stereotyping of domains.

Most of the studies dealing with the gender stereotyping of different academic domains have focused on the masculine stereotyping of the so-called *STEM subjects* (i.e., Science, Technology, Engineering, and Mathematics), which is considered an important factor in explaining the large underrepresentation of women in these fields. When adolescents around the age of 15 were asked whether they conceived of these subjects as more appropriate



for one gender only (using explicit measures), most of them endorsed clear-cut math-male stereotypes and ascribed more talent, ability, and interest in math to boys than to girls (e.g. in France: Chatard, Guimond, & Selimbegovic, 2007, in Germany: Steffens et al., 2010; Steffens & Jelenec, 2011). Similar results were found for Chilean pre-school children (Río & Strasser, 2013). Children in primary school in both France and the U.S., however, perceived girls and boys of their own age as equally talented in math (e.g. Ambady, Shih, Kim, & Pittinsky, 2001) but thought of math as more appropriate for male than female adults (e.g. Martinot, Bagès, & Désert, 2012).

A comprehensive overview of research into gender stereotyping of the *verbal domain* (“reading”) is provided by Sokal (2010). In kindergarten and primary school, the gender neutral option was most often chosen by children if they were asked to classify different kinds of activities (such as reading) as masculine, feminine, or suitable for both genders (e.g. Sokal, Katz, Chaszewski, & Wojcik, 2007 for Canada). Other studies with adolescents applied rating scales to evaluate the perceived femaleness of languages. When the gender connotations of languages were compared with those of math, languages were stereotyped as a female domain in Canada (Plante, Théorêt, & Favreau, 2009) and Germany (Steffens et al., 2010; Steffens & Jelenec, 2011).

Studies measuring *implicit* stereotyping of subjects, however, have found that, on average, school pupils and university students clearly conceive of math and science as male subjects and languages as female subjects. The prevalence and strength of implicit stereotypes have usually been measured with the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998), a computer based measure for assessing the strength of automatic associations between two pairs of concepts, e.g. math and male, language and female. When interpreting the effects, most studies focus exclusively on the implicit gender connotations of a STEM subject while they, in fact, simultaneously inform us about gender associations of the contrasting subject (usually language or arts). Implicit math-male stereotypes have been found

in contrast with art-female (Nosek, Banaji, & Greenwald, 2002; Nosek & Smyth, 2011), language-female (Smeding, 2012; Steffens et al., 2010), verbal-female (Nosek & Smyth, 2011) and reading-female stereotypes (Cvencek, Meltzoff, & Greenwald, 2011) for different age groups ranging from 4-year old U.S. Americans (Cvencek et al., 2011) and German secondary school students (Steffens et al., 2010) to U.S. American (Nosek et al., 2002) and French college students (Smeding, 2012). Strong implicit math-male stereotypes have recently been found in an age- and ethnic diverse sample of more than 5000 adults with various educational backgrounds from different Western, industrialized countries (Nosek & Smyth, 2011).

The relevance of gender stereotyping of domains for students' motivation and performance can be traced back to Kohlberg's (1966) assumption that a person will be more likely to engage in a (learning) activity that fits his or her gender - and to do so more successfully. In line with this, the degree of fit between one's own gender and the perceived maleness or femaleness of academic activities, topics, and domains has been proven an important predictor of academic achievement along with various achievement-related constructs such as attitudes, motivation, and choices. Interestingly, *implicit* domain-gender stereotypes are more powerful in predicting and explaining academic gender differences than explicit domain-gender stereotypes (e.g. Kiefer & Sekaquaptewa, 2007; Lane, Goh, & Driver-Linn, 2012; Nosek & Smyth, 2011; Nosek et al., 2002; Steffens et al., 2010). Implicit math-male stereotypes have been found to be predictive of undergraduate students' as well as adolescents' math performance in Germany and the U.S. (e.g. Kiefer & Sekaquaptewa, 2007; Nosek et al., 2002; Nosek & Smyth, 2011; Steffens & Jelenec, 2011; Steffens et al., 2010). Similar results have been found for emotional and motivational aspects. Implicit domain-gender stereotypes predict less liking (Kessels, Rau, & Hannover, 2006; Nosek et al., 2002), more negative attitudes (Nosek & Smyth, 2011) and lower competence beliefs (Nosek & Smyth, 2011; Steffens & Jelenec, 2011; Steffens et al., 2010) for German and U.S. American

adolescents and adults of the gender stereotyped as incompatible with the academic domain. Implicit domain-gender stereotypes also partially explain gender differences in career preferences (Lane et al., 2012; Nosek & Smyth, 2011; Steffens et al, 2010).

In summary, the gender stereotyping of academic domains is a well-researched field of study. Findings from IAT studies provide robust evidence that math is stereotyped as masculine while languages are stereotyped as feminine. At the same time, studies show how important gender-appropriateness is for liking, interest, and performance in particular domains. In line with this, it is to be expected that, *if* learning and academic engagement might be stereotyped as feminine, the detrimental effects on boys' academic engagement should be most visible in German (as a feminine domain) and less pronounced in math (as a masculine domain).

### **Gender Stereotyping of School and Learning**

As compared with the large body of research on the gender stereotyping of domains, surprisingly few studies exist that *directly* measure general gender stereotyping of school and learning activities, and those are quite dated. Those studies (Kagan, 1964; Kellogg, 1969) asked U.S. American primary school children of different grades to classify pictures of school related objects as more suitable for either male or female use, yielding inconsistent results: On the one hand, grade 2 boys - but not grade 3 boys- more often perceived book and blackboard as feminine than masculine (Kagan, 1964). On the other hand, boys and girls more often considered book, blackboard, library, and chalk as feminine, but map and pencil as masculine (Kellogg, 1969). Interestingly, the boys and girls in both studies perceived both school and desk as more appropriate for their *own* gender, which suggests that primary school children associate being at school mainly with themselves. Taken together, these few studies show little direct empirical evidence for a clear-cut gendered connotation of school, but possible feminine connotations of school and learning could also be derived more indirectly from

empirical evidence that shows the school setting is perceived as suiting femininity better than masculinity.

**School as a setting rewarding femininity.** The feminisation of school has been discussed as one possible explanation for boys' underachievement. The term feminisation is used in different ways. It can refer to the (factual) overrepresentation of women in the teaching profession. Despite widespread circulation of the claim in the media that boys' underachievement is due to female teachers (e.g. "Female teachers accused of giving boys lower marks", Garner, 2012), empirical evidence supporting this view is scarce. Many studies from different countries have provided no evidence for the assumption that boys achieve more highly or are more motivated if taught by a male teacher ( e.g. Carrington, Tymms, & Merell, 2008; Martin & Marsh, 2005; Neugebauer, Helbig, & Landmann, 2011).

But feminisation of school can also mean that the school setting itself is feminine rather than masculine, in the sense that behaviour congruent with the female gender role is expected and rewarded, while behaviour congruent with the male gender role is sanctioned. In line with this, some authors have linked the difficulties boys experience in school to the "incompatible demands of the male sex role and student role" (Gold & Reis, 1982; p. 508), and have argued that the male gender role conflicts with academic engagement at school (e.g. Mickelson, 1989). This view is supported by findings from a Canadian study which revealed that primary school teachers liked 6- and 7-year old boys they perceived as being masculine less than boys of the same age they perceived as feminine or androgynous and also saw them as displaying more problematic externalizing and less prosocial behaviour (Piché & Plante, 1991). Orr (2011) argues that boys are socialized to display specific masculine characteristics (such as being independent, assertive, and competitive), which might at present cause problems in school. In contrast, girls' socialization tends to encourage more dependent, cooperative, social individuals whose habitus is more consistent with the school environment. In a nationally representative sample of more than 6000 children from the U.S. at the age of 5

and 6, Orr (2011) tested these assumptions and found that girls were engaged more often than boys in female activities (such as singing, chores) and had more positive school attitudes and better grades, whereas boys were more engaged in male activities (such as building, sports) and had more negative school attitudes and lower grades than girls. More importantly, the more parents made their child (of both genders) participate in female activities, the better the child's grades in kindergarten were.

It seems premature, however, to conclude that either the male gender role itself or the mere possession of masculine traits should lead to lower academic performance. A great deal of research has even shown the beneficial effects of agency or instrumentality on indicators of academic or vocational success. For instance, Vollmer (1986) found that instrumentality was correlated with perceived ability in a sample of psychology students. In addition, instrumentality was found to be positively related to U.S. American undergraduates' problem-solving skills (Heppner, Walther, & Good, 1995) and to Swiss adolescents' career-related goal clarity (Freund, Weiss, & Wiese, 2013). For a better understanding of *which* specific aspects of masculinity might clash with a possibly feminine school setting, it is necessary to disentangle aspects of masculinity that represent socially valued, positive traits, such as being active, independent and self-confident, from other aspects of masculinity that are more likely to cause problems at school. This has been achieved through theoretical conceptualisations of the gender role self-concept that include both positive and negative traits (e.g. Krahe, Berger, & Möller, 2007; Runge, Frey, Gollwitzer, Helmreich, & Spence, 1981; Spence, Helmreich, & Holahan, 1979). Research building on these approaches measures socially desirable (positive) masculine traits and socially undesirable (negative) masculine traits on separate scales. Using this differentiation, a German study (Kessels & Steinmayr, 2013a) has revealed that only *negative* masculinity is significantly (negatively) associated with both school performance and academic help-seeking attitudes in a sample of 11th-graders.

Taken together, studies related to the possible feminisation of school suggest that the school setting rewards feminine traits and behaviour more than it does masculine traits and behaviour, as children high on masculinity achieve worse grades and are less liked by teachers. In addition, it has become clear that not all aspects of a masculine self-concept should conflict with academic demands, mainly those comprised of socially undesirable masculine traits.

**Academic disengagement as a way to establish masculinity.** A different strand of research studying a possible misfit between boys' masculine self-concept and their perception of school works primarily with ethnographic, qualitative measures such as observations, interviews, and group discussions. For instance, in extending Epstein's (1998) work on *failing boys*, Jackson and colleagues (Jackson, 2002; Jackson & Dempster, 2009) interviewed sixth graders and university students in the U.K. on their gendered perception of academic work. The interviews revealed that if male students wanted to be perceived as cool, as masculine, and as *naturally* able (Jackson & Dempster, 2009; p.342), they generally had to avoid overt displays of hard work. Morris (2008) interviewed students from a U.S. American rural high school and concluded that lack of academic effort was an important aspect of boys' masculine identity. Similarly, Mac an Ghail (1994) described the way male academic achievers were positioned by their classmates as *effeminate* and consequently bullied because of their perceived masculine *lack*. The *lads*, the *real hard* or *cool* boys, in contrast, performed their masculinity by demonstrating an explicit anti-school ethos and distancing themselves from academic achievement, authorities, and following rules in school in England (e.g. Francis, 1999; Jackson, 2003; Jackson & Dempster, 2009; Willis, 1977) and Australia (Connell, 1989). Very similar to the findings by Kessels and Steinmayr (2013a) that only *negative* masculinity has an impact on academic achievement, these qualitative studies also suggest that only a *specific kind of masculinity*, related to terms like "laddishness", is supported by distance from school.

Taken as a whole, the discussion on boys' underachievement at school often suggests a lack of fit between a feminine school setting and the masculine self-concept of boys. Claims of feminine connotations for academic engagement, however, have been, for the most part, negatively derived from descriptions of the way boys strengthen their masculinity by opposing school. To conclude from these findings that school in general is actually perceived as feminine by the majority of students seems premature. A more direct and up-to-date assessment is needed. In addition, extant research suggests socially desirable masculinity (as related to instrumentality) should not cause boys' academic problems, but socially undesirable masculine traits related to so-called laddishness might.

### **Study Overview and Hypotheses**

The aim of this study is to test to what extent school in general has acquired a feminine connotation and whether this potentially feminine connotation has an impact on boys' achievement. In line with recent studies measuring gender stereotyping of academic domains, we used an implicit measure to assess the gender stereotyping of school. We expected that students would associate school more strongly with female than with male (*Hypothesis 1*). We used the Go/No-go Association Task (GNAT, Nosek & Banaji, 2001), which examines automatic cognition toward a *single* target category without requiring the direct involvement of complementary or contrasting objects as the IAT does. In this way we circumvented methodological and logical problems which might arise from gender stereotyping of potential complementary categories (cf. Blanton, Jaccard, Gonzales, & Christie, 2006; Fiedler, Messner, & Bluemke, 2006).

As the recent discussion on boys' underachievement at school suggests that a lack of fit between a feminine school and a masculine self-concept in boys causes boys' lower achievement (e.g. Helbig, 2010; Jackson & Dempster, 2009; Mickelson, 1989), we expected that an implicit feminine stereotyping of school is especially detrimental to boys' achievement

if they regard themselves as highly masculine. Because earlier studies (e.g. Kessels & Steinmayr, 2013a) have revealed that *negative* masculinity is more strongly linked to boys' disengagement from school than is positive masculinity, we used two separate scales measuring socially desirable (positive) and socially undesirable (negative) masculine traits (cf. Krahe et al., 2007; Runge et al., 1981; Spence et al., 1979).

As different academic domains are gender stereotyped (e.g. Nosek et al., 2002; Steffens et al., 2010), we expected that the detrimental impact on boys' achievement of the school-female stereotype would differ in different school subjects: It should be most pronounced in German, because the verbal domain is stereotyped as feminine (e.g. Plante et al., 2009; Steffens et al., 2010), making high achievement in this domain least compatible with boys' perception of themselves as highly masculine and engagement at school as something feminine. Achievement in math should not be impaired by boys' general perception of school as feminine, as math is stereotyped as masculine (e.g. Chatard et al., 2007; Nosek et al., 2002), and good grades should not threaten boys' self-concept, even if they perceive of themselves as highly masculine and school in general as something feminine. In summary, we expected that the more strongly boys associated school with female (rather than male) and the more they ascribed negative masculine traits to themselves, the lower would be their grades in German (*Hypothesis 2a*), while for math, this relationship should be less pronounced (*Hypothesis 2b*).

## Method

### Participants

One hundred twenty-three 9th-grade students of a comprehensive school located in a large city in the western part of Germany participated in this study. In nationwide standardized achievement assessment, the school's environment was categorized as "difficult/



challenging", meaning there is relatively higher number of students of minority background and/or from educationally deprived families. One female participant was excluded due to excessive error rates (> 30% in both critical blocks), likely indicating poor understanding of the task. The remaining sample consisted of  $N=122$  students, 65 female, 57 male. Their average age was  $M = 14.45$  ( $SD = 0.60$ ). 41 of them (34%) reported German as not being their mother tongue. The sample is characterized in more detail in Table 1. The students participated voluntarily during regular school hours and received a chocolate bar for participation.

Insert Table 1 about here

## Materials

**Go/No-go Association Task (GNAT).** The GNAT assesses the strength of implicit association between a single target concept (school) and two attribute categories (female, male). In the GNAT procedure, stimuli from the target concept and from one attribute category serve as signal items and require pressing the space bar as a correct response (go trials). Stimuli from the other attribute category serve as distractors and participants are instructed not to react to them (no-go trials). The higher the accuracy in discriminating signal items from noise items, the stronger the implicit association between the target concept and the attribute concept that serve as signal items.

Our study examined the strength of association between the target concept school (stimuli: taking an active part in lessons, paying attention, studying) and the attribute categories male (stimuli: man, boy, son) and female (stimuli: woman, girl, daughter). The selected stimuli representing the target concept were pretested with  $N = 47$  students as being terms highly associated with school. The stimuli for representing male and female were not pretested due to their obvious gender but were selected based on theoretical considerations such as familiarity and equal valence.

The GNAT consisted of five blocks differing in the categories serving as signal items. Within each block, the names of the categories representing signal items were visible in the upper part of the screen as a task reminder. The stimuli appeared consecutively and in random order in the middle of the screen. The first and second blocks were practice blocks, in which stimuli from either the category female or the category male were signal items requiring pressing the space-bar as a correct response. The first block was selected randomly. The third block was always the block for practicing the target category school (signal items: school, distractor items: male, female). Each of the three practice blocks consisted of 12 trials, half of them representing signal items, half of them distractors. Blocks 4 and 5 provided critical data by pairing school with either female or male. Here, stimuli from two categories served simultaneously as signal items and required pressing the space bar as a correct response. Each of the blocks consisted of 60 trials, half of them representing signal items, half of them distractors. The order of stimuli within the block and the order of pairing in the critical blocks was selected randomly. The response window for each trial was 2000ms in the practice blocks to facilitate understanding of the task and 1000ms in the critical blocks. After each trial, a green check marking correct responses or a red f (abbreviation of the German “falsch”) marking responses that were wrong or too slow appeared for 500ms giving participants feedback on their reaction. The error rates were 6% in the school-female block and 8% in the school-male block.

The sensitivity index  $d'$  indicates the strength of association between the two concepts that served as signal items (cf. signal detection theory, e.g. Green & Swets, 1966). The proportions of hits (correct ‘go’ response on signal items) and false alarms (incorrect ‘go’ responses on distractors) were converted to z-scores and subtracted from one another resulting in  $d'$ : the larger  $d'$ , the stronger the association between the concepts.  $d'$  was computed for each critical block separately. The difference between  $d'$  *school-female* and  $d'$  *school-male* within each participant showed whether he or she associated school more strongly with

female than with male or vice versa. Positive  $d'$ -difference scores indicate a stronger association with female, negative  $d'$ -difference scores a stronger association with male. Empty cells, which result in infinity when applying the  $z$  function, were corrected in a model-sensitive way (Banaji & Greenwald, 1995).

**Socially desirable and undesirable masculinity.** Socially desirable and undesirable masculine traits (M+ and M-) were measured explicitly by the corresponding subscales of the German version of the Extended Personal Attribute Questionnaire (Spence et al., 1979). The German version is a validated translation with psychometric properties similar to the U.S. questionnaire (Runge et al., 1981). The M+ scale consisted of seven items (e.g. self-confident, active) and the M- scale of nine items (e.g. arrogant, aggressive; see Appendix). All items were rated on a 6-point scale (e.g. (1) *not at all aggressive*, (6) *very aggressive*). In order to facilitate comparing M+ and M- despite their unequal number of items, we computed means rather than sums of the scale item scores. The reliabilities were  $\alpha = .62$  for M+ and  $\alpha = .65$  for M-. In order to enhance the reliability of M+, we omitted the item “independent”, which improved the reliability of the scale the most, resulting in  $\alpha = .68$ . (The omission of this item did not alter the results of the following analyses.) Although the reliabilities obtained are still rather low, they are comparable with those reported in earlier studies using the same scales for a similar age group (Athenstaedt, Mikula, & Brecht, 2009).

**School achievement.** Previous report card grade and course level (basic course or advanced course, resulting from tracking within the comprehensive school) for the subjects German and math were reported by the participants. In Germany, grades range from 1 (*outstanding performance*) to 6 (*insufficient performance*). For comparing grades between different course levels, the grades in the basic course were recoded by adding 1 to each grade, resulting in a grade range from 1 to 7.

**Demographics.** Participants indicated gender, age, and mother tongue.

## Procedure

Students were tested in groups of eight by two experimenters. First, the GNAT was administered on portable netbooks using DirectRT software (Jarvis, 2010). After performing the GNAT, participants completed the paper-and-pencil questionnaire including the measures described above. Finally, they were rewarded with a bar of chocolate and debriefed.

## Results

Means of all variables used in the analyses are depicted separately for boys and girls in Table 2, bivariate correlations separately for boys and girls in Table 3. The correlation matrix showed that socially desirable masculinity did not correlate with any of the variables to be used for testing our research hypotheses, so it was excluded from further analyses (cf. Kessels & Steinmayr, 2103a). In order to test for overall gender differences, a MANOVA with student's gender as a factor was conducted on all variables included in the analyses. It showed a multivariate main effect of gender,  $F(6,115) = 3.90, p = .001, \text{Wilks' } \lambda = .83, \text{partial } \eta^2 = .17$ . Univariate tests revealed the following statistically significant gender differences: Boys had significantly higher values in socially desirable masculinity but not in undesirable masculinity (see Table 2). In addition, girls associated school more strongly with female than boys did, and more strongly with female as compared to male than boys did. With regard to academic performance, boys had lower grades than girls in German, but no gender differences occurred in math grades.

Insert Table 2 and Table 3 about here

## Implicit Stereotyping of School

*Hypothesis 1* stated that participants would implicitly associate school more with female than with male. For the whole sample, the  $d'$ -difference score ( $M = 0.20, SD = 1.29$ ) was found to be significantly larger than zero, which indicates a stronger implicit association of school with female than with male,  $t(121) = 1.69, p = .048$  one-tailed. Overall, this

corroborates Hypothesis 1. However, this result was exclusively due to the girls' performance: Girls demonstrated higher accuracy in the school-female block than in the school-male block,  $t(64) = 2.92, p = .005$ , whereas the boys' performance did not differ between the two blocks,  $t(56) = -0.45, p = .654$ .

### **Effects of Implicit Stereotyping on Academic Achievement**

In *Hypothesis 2* we assumed that the more boys associated school with female than with male and the more they ascribed negative masculine traits to themselves, the lower would be their grades in German (*Hypothesis 2a*), while for math, this relationship should be less pronounced (*Hypothesis 2b*). In order to test the hypothesized relationships, we performed two-step hierarchical regression analyses in each subject, separately for boys and girls. To avoid multicollinearity, all continuous variables were centered before being included in the regression equation. After that, the Variance Inflation Factors (VIFs) of all predictors were  $\leq 1.12$  indicating a low degree of multicollinearity (for a critical discussion of the rules of thumb for interpreting VIFs cf. e.g. O'Brien, 2007).

**German grades.** Analysis of the male subsample: In the first step, boys' German grades were regressed on socially undesirable masculinity and  $d'$ -difference. In the second step, their interaction was included (see Table 4). The model was statistically significant overall in step 1,  $F(2,54) = 5.86, p = .005, R^2 = .18$ , and step 2,  $F(3,53) = 5.49, p = .002, R^2 = .24$ . In the first step, socially undesirable masculinity predicted German grades, in that the more socially undesirable masculine traits boys ascribed to themselves, the worse their German grades were,  $\beta = .40, t(54) = 3.17, p = .002$ . Socially undesirable masculinity remained a statistically significant predictor after including the interaction term in step 2. The statistically significant interaction between socially undesirable masculinity and the implicit association of school,  $\beta = .25, t(53) = 2.02, p = .048$ , indicated that the more school was implicitly associated with female than with male and the more socially undesirably masculine traits the boys ascribed to themselves, the worse their German grades were (see Figure 1).

Simple slope analyses (Aiken & West, 1991) showed a statistically significant slope for boys with high scores in socially undesirable masculinity, i.e. 1 *SD* above the mean,  $\beta = .45$ ,  $t(53) = 2.69$ ,  $p = .01$ , but not for boys with low values in socially undesirable masculinity, i.e. 1 *SD* below the mean,  $\beta = -.02$ ,  $t(53) = -0.11$ ,  $p = .909$ .

Analysis of the female subsample: For girls, however, the overall model did not reach statistical significance (see Table 4). Neither socially undesirable masculinity nor the implicit association of school nor their interaction predicted girls' grades in German to a statistically significant degree.

Insert Table 4 and Figure 1 about here

**Math grades.** Analysis of the male subsample: Boys' math grades were regressed on socially undesirable masculinity and *d'*-difference in the first step, and in the second step their interaction was included. The overall model did not reach statistical significance (see Table 4). Socially undesirable masculinity was found to be a marginally significant predictor in step 1, but not in step 2. The implicit gender stereotyping of school and the interaction effect did not predict boys' grades in math to a statistically significant degree.

Analysis of the female subsample: For girls, as well, the overall model did not reach statistical significance (see Table 4). An unexpected interaction effect emerged, however,  $\beta = .28$ ,  $t(61) = 2.18$ ,  $p = .033$ . Closer examination of the interaction revealed that the more girls associated school with male than with female and the more socially undesirable masculine traits they ascribed to themselves, the better their grades in math were (see Figure 2).

However, simple slope analyses (Aiken & West, 1991) showed no statistically significant slope for girls with high scores in socially undesirable masculinity, i.e. 1 *SD* above the mean,  $\beta = .29$ ,  $t(61) = 1.64$ ,  $p = .107$ , or for girls with low values in socially undesirable masculinity, i.e. 1 *SD* below the mean,  $\beta = -.15$ ,  $t(61) = -0.84$ ,  $p = .403$ . Because neither the overall test nor the simple slopes reached statistical significance, and no theoretically driven hypothesis predicted this effect, we treat this finding with caution. We cannot exclude the possibility that

the interaction term reached statistical significance in the regression analysis only because of the problems associated with multiple testing.

Insert Figure 2 about here

### Discussion

This study used an implicit measure, the GNAT, to investigate whether school in general has acquired a feminine connotation and whether this potentially feminine connotation has a detrimental impact on boys' academic achievement. The GNAT revealed an implicit school-female association for girls only, but no gender stereotyping of school for boys, a finding that only partially supports our Hypothesis 1. School=female seems *not* to be a stereotype automatically activated when students think about school, as it is in the case of math=male or languages=female. This indicates that school is not generally perceived as feminine, as has been suggested in the discussion concerning the feminisation of school, even though most teachers are female. The understanding of the GNAT as a measure of individual differences (cf. Lane et al., 2012; Nosek et al., 2010; Smeding, 2012) implies, however, that there is a subgroup of boys who also perceive school as feminine. We expected a school-female association for boys to be more detrimental in the female-stereotyped subject German than in the male-stereotyped subject math and most detrimental for boys with a masculine self-concept, because they should perceive a greater misfit between their own gender and school. As predicted by Hypothesis 2a, we found that the stronger the school-feminine association was and the more boys described themselves with socially undesirable masculine traits, the lower their grades in German were. In line with Hypothesis 2b, no effects on grades in math were found.

Overall, our findings indicate that the majority of male students do not link school with female on the level of implicit automatic associations. Those boys, however, who do not follow the dominant pattern of associating school with their own gender, actually do achieve

less in the feminine stereotyped domain of languages. This is particularly the case for those displaying “macho-like”-traits. In line with previous qualitative research on “hard” or “laddish” boys distancing themselves from academic engagement (cf. Francis, 1999; Jackson, 2002; Jackson & Dempster, 2009; Willis, 1977), we found that only a specific kind of masculinity characterized by undesirable masculine traits, combined with school-female stereotypes, is at odds with academic achievement. Importantly, the differential effect of negative masculinity on boys and girls is not due to an absolute difference in negative masculinity, as boys and girls did not differ in self-reported undesirable masculine traits. A recent study of Austrian adolescents in a similar age group (Athenstaedt et al., 2009) also found no gender differences on negative masculinity as measured with the GEPAQ. This is in line with recent findings from Germany that stereotypes of women and their gender role are becoming more flexible, as women are perceived to be increasing their masculine characteristics while maintaining their feminine ones, while stereotypes of men have been found to be relatively stable in Germany (Wilde & Dieckman, 2005).

The gender-specific interaction terms in our regression analyses reveal that it is not a single variable (such as an implicit gender connotation of school or a highly masculine self-concept) but the interplay between one’s own gender, gender role self-concept, and gender stereotyping of academic domains and school in general which is crucial in explaining gender differences in academic achievement. Overall, our results point out the importance of fit between gender role self-concept and task characteristics, supporting extant findings on the relevance of gender appropriateness for liking particular subjects (e.g. Kessels et al., 2006), competence beliefs (e.g. Steffens et al., 2010), and academic performance (e.g. Nosek et al., 2002). Furthermore, our findings can be embedded in the large and diverse body of research on the relationship between identity and academic interests and motivation. Osborne and Jones (2011), for instance, have elaborated theoretically on the way in which the structure of the self influences academic outcomes: Adolescents who strongly define their selves through



their role as students are more likely to be motivated and successful in school than academically unidentified students. Similarly, academic engagement increases when an academic task is perceived as identity-congruent, for example, when congruent with gender identity (e.g. Elmore & Oyserman, 2012; Kessels & Hannover, 2007; Oyserman & Destin, 2010).

Limitations of our study concern the correlational and cross-sectional nature of our data which do not allow for causal inferences. Theoretically, our hypothesis could also be formulated from the opposite direction: If a boy shows low performance in German, he perceives school as something feminine. This direction would be plausible given the present emphasis on verbal skills in nearly *all* school subjects in the school curricula in Germany. Following this line of argumentation, Kessels and Steinmayr (2013b) found that only students' verbal self-concept predicted valuing school in general, whereas the math self-concept was unrelated to valuing school. Applied to our results, boys perceiving a poor fit between themselves and the verbal domain might generalize this perception to a misfit between themselves and school in general, possibly leading them to disengage from learning. Because we did not assess the gender stereotyping of German and math in our sample, we could not test whether the boys who are high on socially undesirable masculine traits and who perceive school as feminine also hold language-female stereotypes. The large amount of evidence seems, however, to permit this assumption (e.g. Cvencek et al., 2011; Plante et al., 2009; Steffens et al., 2010). Beyond that, our study is limited to a German sample of ninth-graders. As gender roles are culturally shaped and develop over time (cf. Twenge, 1997; Wilde & Dieckman, 2005), different correlations might be observed for a different age or cultural group. Moreover, our study focussed entirely on a specific set of gender-related variables and did not include additional structural variables, such as ethnicity or socioeconomic status (SES), which are predictive of academic achievement and may interact

with the understanding and construction of masculinity and femininity in the school context (cf. Legewie & DiPrete, 2012).

For boys who describe themselves as laddish-like and perceive school as feminine, our study has highlighted the relevance for academic achievement of the fit between the characteristics of the learner and the environment. Practical implications suggest strategies to enhance this fit by addressing the gendered connotations of individual subjects and school in general as well as boys' situational awareness of their gender. In line with existing studies (e.g. Dasgupta & Asgari, 2004), exposure to counterstereotypic persons may reduce the gender connotations of academic subjects. Cross-sectionally, boys with male German teachers perceived German as less explicitly feminine than boys with female German teachers (Hadjar & Lupatsch, 2011). Promoting role models of both sexes who represent high achievement in all areas of study at school may, therefore, be one strategy to reduce the gender stereotyping of subjects. Beyond that, reducing the salience of gender in school may be a way to make student's self-concept less relevant and thus increase engagement in those subjects or activities deemed inappropriate to one's own gender (cf. Elmore & Oyserman, 2012). For instance, research has shown that girls are more engaged in single-gender physics classes than in mixed-gender physics classes, because in single gender groups gender-related self-knowledge is less accessible (Kessels & Hannover, 2008).

Future research might look for factors explaining individual differences in the gender stereotyping of school and for institutional factors which have an impact on the gender stereotyping of school. Dresel, Stöger, and Ziegler (2006) have shown that the gender gap in educational achievement strongly varies between schools, a finding that might be due to differences between schools in their gender stereotyping of learning. Legewie and DiPrete (2012) explain in more detail how institutional factors might moderate the gender gap in achievement. They argue that school environments - via different peer group norms - shape specific conceptions of masculinity, which will either foster or inhibit boys' anti-school ethos.

Specifically, they have proposed that in classes with higher SES, academic competition is promoted as an aspect of masculine identity, while in classes with lower SES, masculinity is more likely to be constructed via oppositional behavior. In support of these hypotheses, they have found for two representative German samples of different age groups that boys benefit academically more than girls from being in a classroom of high SES. It remains a task for future research to develop and test interventions which may enhance the perceived fit between self-concepts and learning for both boys *and* girls of all SES backgrounds.

### References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385–390. doi:10.1111/1467-9280.00371
- Athenstaedt, U., Mikula, G., & Bredt, C. (2009). Gender role self-concept and leisure activities of adolescents. *Sex Roles, 60*, 399–409. doi:10.1007/s11199-008-9543-y
- Banaji, M. R., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology, 68*, 181–198. doi:10.1037/0022-3514.68.2.181
- Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42*, 192–212. doi:10.1016/j.jesp.2005.07.003
- Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*, 319–337. doi:10.1146/annurev.soc.34.040507.134719
- 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. doi:10.1080/01411920701532202
- 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. doi:10.1016/j.jesp.2006.10.024
- Clark, M. A., Lee, S. M., Goodman, W., & Yacco, S. (2008). Examining male underachievement in public education: Action research at a district level. *NASSP Bulletin, 92*, 111–132. doi:10.1177/0192636508321155
- Connell, R. W. (1989). Cool guys, swots and wimps: The interplay of masculinity and education. *Oxford Review of Education, 15*, 291–203. doi:10.1080/0305498890150309
- Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development, 82*, 766–779. doi:10.1111/j.1467-8624.2010.01529.x
- Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642–658. doi:10.1016/j.jesp.2004.02.003
- Dresel, M., Stöger, H., & Ziegler, A. (2006). Klassen- und Schulunterschiede im Ausmaß von Geschlechterdiskrepanzen bei Leistungsbewertungen und Leistungsaspirationen: Ergebnisse einer Mehrebenenanalyse [Class and school differences in gender discrepancies in performance evaluations and achievement aspirations: Results of a multilevel analysis]. *Psychologie in Erziehung und Unterricht, 53*, 44–61.
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198–208. doi:10.1037/0022-0663.98.1.198
- Elmore, K. C., & Oyserman, D. (2012). If ‘we’ can succeed, ‘I’ can too: Identity-based motivation and gender in the classroom. *Contemporary Educational Psychology, 37*, 176–185. doi:10.1016/j.cedpsych.2011.05.003

- Epstein, D. (1998). Real boys don't work: 'underachievement', masculinity and the harassment of 'sissies'. In D. Epstein, J. Elwood, V. Hey, & J. Maw (Eds.), *Failing boys? Issues in gender and achievement* (pp. 96–108). Buckingham: Open University Press.
- Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the “I”, the “A”, and the “T”: A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*, 74–147. doi:10.1080/10463280600681248
- Francis, B. (1999). Lads, lasses and (new) labour: 14-16-year-old students' responses to the 'laddish behaviour and boys' underachievement' debate. *British Journal of Sociology of Education*, *20*, 355–371. doi:10.1080/01425699995317
- Freund, A. M., Weiss, D., & Wiese, B. S. (2013). Graduating from high school: The role of gender-related attitudes, self-concept and goal clarity in a major transition in late adolescence. Advance online publication. *European Journal of Developmental Psychology*. doi:10.1080/17405629.2013.772508
- Garner, R. (2012, February 16). Female teachers accused of giving boys lower marks. *The Independent*. Retrieved from <http://www.independent.co.uk/news/education/education-news/female-teachers-accused-of-giving-boys-lower-marks-6943937.html>
- Gold, D., & Reis, M. (1982). Male teacher effects on young children: A theoretical and empirical consideration. *Sex Roles*, *8*, 493–513. doi:10.1007/BF00287715
- Gray, J., Peng, W.-J., Steward, S., & Thomas, S. (2004). Towards a typology of gender-related school effects: Some new perspectives on a familiar problem. *Oxford Review of Education*, *30*, 529–550. doi:10.1080/0305498042000303991
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi:10.1037/0022-3514.74.6.1464
- Hadjar, A., & Lupatsch, J. (2011). Geschlechterunterschiede im Schulerfolg: Spielt die Lehrperson eine Rolle? [Gender differences in educational success: Does the teacher matter?]. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *31*, 79–94.
- Hannover, B., & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze [Are boys left behind? Reviewing and explaining education-related gender disparities]. *Zeitschrift für Pädagogische Psychologie*, *25*, 89–103. doi:10.1024/1010-0652/a000039
- Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? [Are female teachers responsible for the school performance gap between boys and girls?]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *62*, 93–111. doi:10.1007/s11577-010-0095-0
- Heppner, P. P., Walther, D., & Good, G. E. (1995). The differential role of instrumentality, expressivity, and social support in predicting problem-solving appraisal in men and women. *Sex Roles*, *32*, 91-108. doi:10.1007/BF01544759
- Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy. *Gender and Education*, *14*, 37–50. doi:10.1080/09540250120098870
- Jackson, C. (2003). Motives for 'laddishness' at school: Fear of failure and fear of the 'feminine'. *British Educational Research Journal*, *29*, 583–598. doi:10.1080/01411920301847

- Jackson, C., & Dempster, S. (2009). 'I sat back on my computer ... with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, 18, 341–356. doi:10.1080/09589230903260019
- Jarvis, B. G. (2010). DirectRT. New York, NY: Empirisoft Corporation.
- Kagan, J. (1964). The child's sex role classification of school objects. *Child Development*, 35, 1051–1056.
- Kellogg, R. L. (1969). A direct approach to sex-role identification of school-related objects. *Psychological Reports*, 24, 839–841. doi:10.2466/pr0.1969.24.3.839
- Kessels, U., & Hannover, B. (2007). How the image of math and science affects the development of academic interest. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report of the DFG priority programme* (pp. 283–297). Münster: Waxmann.
- Kessels, U., & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. *British Journal of Educational Psychology*, 78, 273–289. doi:10.1348/000709907X215938
- Kessels, U., Rau, M., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 76, 761–780. doi:10.1348/000709905X59961
- Kessels, U., & Steinmayr, R. (2013a). 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. doi:10.1016/j.lindif.2012.09.013
- Kessels, U., & Steinmayr, R. (2013b). Der subjektive Wert von Schule in Abhängigkeit vom verbalen und mathematischen Selbstkonzept [Valuing school as a function of verbal and mathematical ability self-concepts]. *Zeitschrift für Pädagogische Psychologie*, 27, 105–113. doi:10.1024/1010-0652/a000093
- Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science*, 18, 13–18. doi:10.1111/j.1467-9280.2007.01841.x
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–172). Stanford, CA: Stanford University Press.
- Köller, O., Daniels, Z., Schnabel, K. U., & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse [Course selections of girls and boys in mathematics: The role of academic self-concept and interest]. *Zeitschrift für Pädagogische Psychologie*, 14, 26–37. doi:10.1024//1010-0652.14.1.26
- Krahé, B., Berger, A., & Möller, I. (2007). Entwicklung und Validierung eines Inventars zur Erfassung des Geschlechtsrollen-Selbstkonzepts im Jugendalter [Development and validation of an inventory for measuring gender role self-concept in adolescence]. *Zeitschrift für Sozialpsychologie*, 38, 195–208. doi:10.1024/0044-3514.38.3.195
- Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles*, 66, 220–234. doi:10.1007/s11199-011-0036-z

- Legewie, J., & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review*, *77*, 463–485. doi:10.1177/0003122412440802
- Lindsay, G., & Muijs, D. (2006). Challenging underachievement in boys. *Educational Research*, *48*, 313–332. doi:10.1080/00131880600992389
- Mac an Ghail, M. (1994). *The making of men*. Buckingham: Open University Press.
- Martin, A. J., & Marsh, H. (2005). Motivating boys and motivating girls: Does teacher gender really make a difference? *Australian Journal of Education*, *49*, 320–334. doi:10.1177/000494410504900308
- Martinot, D., Bagès, C., & Désert, M. (2012). French children's awareness of gender stereotypes about mathematics and reading: When girls improve their reputation in math. *Sex Roles*, *66*, 210–219. doi:10.1007/s11199-011-0032-3
- Mickelson, R. A. (1989). Why does Jane read and write so well? The anomaly of women's achievement. *Sociology of Education*, *62*, 47–63.
- Morris, E. W. (2008). "Rednecks," "rutters," and arithmetic: Social class, masculinity, and schooling in a rural context. *Gender & Society*, *22*, 728–751. doi:10.1177/0891243208325163
- National Center for Education Statistics. (2004). *Trends in educational equity of girls & women: 2004*. Retrieved from <http://nces.ed.gov/pubs2005/2005016.pdf>
- Neugebauer, M., Helbig, M., & Landmann, A. (2011). Unmasking the myth of the same-sex teacher advantage. *European Sociological Review*, *27*, 669–689. doi:10.1093/esr/jcq038
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*, 625–666. doi:10.1521/soco.19.6.625.20886
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59. doi:10.1037/0022-3514.83.1.44
- Nosek, B. A., & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal*, *48*, 1125–1156. doi:10.3102/0002831211410683
- O'Brien, R. M. (2007). A caution regarding rules of thumb for Variance Inflation Factors. *Quality & Quantity*, *41*, 673–690. doi:10.1007/s11135-006-9018-6
- OECD. (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school*. Paris: OECD.
- Orr, A. J. (2011). Gendered capital: Childhood socialization and the “boy crisis” in education. *Sex Roles*, *65*, 271–284. doi:10.1007/s11199-011-0016-3
- Osborne, J. W., & Jones, B. D. (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. doi:10.1007/s10648-011-9151-1
- Oyserman, D., & Destin, M. (2010). Identity-Based Motivation: Implications for intervention. *The Counseling Psychologist*, *38*, 1001–1043. doi:10.1177/0011000010374775
- Piché, C., & Plante, C. (1991). Perceived masculinity, femininity and androgyny among primary school boys: Relationships with the adaptation level of these students and the attitudes of the teachers towards them. *European Journal of Psychology of Education*, *6*, 423–435. doi:10.1007/BF03172775

- Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology, 29*, 385–405. doi:10.1080/01443410902971500
- Río, M. F., & Strasser, K. (2013). Preschool children's beliefs about gender differences in academic skills. *Sex Roles, 68*, 231–238. doi:10.1007/s11199-012-0195-6
- Runge, T. E., Frey, D., Gollwitzer, P. M., Helmreich, R. L., & Spence, J. T. (1981). Masculine (instrumental) and feminine (expressive) traits: A comparison between students in the United States and West Germany. *Journal of Cross-Cultural Psychology, 12*, 142–162. doi:10.1177/0022022181122002
- Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles, 67*, 617–629. doi:10.1007/s11199-012-0209-4
- Sokal, L. (2010). Prevalence of gendered views of reading in Thailand and Canada. *Alberta Journal of Educational Research, 56*, 44–56.
- Sokal, L., Katz, H., Chaszewski, L., & Wojcik, C. (2007). Good-bye, Mr. Chips: Male teacher shortages and boys' reading achievement. *Sex Roles, 56*, 651–659. doi:10.1007/s11199-007-9206-4
- Spence, J. T., Helmreich, R. L., & Holahan, C. K. (1979). Negative and positive components of psychological masculinity and femininity and their relationships to self-reports of neurotic and acting out behaviors. *Journal of Personality and Social Psychology, 37*, 1673–1682. doi:10.1037/0022-3514.37.10.1673
- Steffens, M. C., & 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. doi:10.1007/s11199-010-9924-x
- Steffens, M. C., 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. doi:10.1037/a0019920
- Tiedemann, J., & Faber, G. (1994). Mädchen und Grundschulmathematik: Ergebnisse einer vierjährigen Längsschnittuntersuchung zu ausgewählten geschlechtsbezogenen Unterschieden in der Leistungsentwicklung [Girls and primary school mathematics: Results from a four-year longitudinal study on selected gender differences of the performance development]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 26*, 101–111.
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles, 36*, 305–325. doi:10.1007/BF02766650
- Vollmer, F. (1984). Sex differences in personality and expectancy. *Sex Roles, 11*, 1121–1139. doi:10.1007/BF00288138
- Wilde, A., & Diekman, A. B. (2005). Cross-cultural similarities and differences in dynamic stereotypes: A comparison between Germany and the United States. *Psychology of Women Quarterly, 29*, 188–196. doi:10.1111/j.1471-6402.2005.00181.x
- Willis, P. E. (1977). *Learning to labour: How working class kids get working class jobs*. New York, NY: Columbia University Press.



## Appendix

*English items of the Extended Personality Questionnaire representing socially desirable and undesirable masculinity and their German translation (Runge et al., 1981).*

Scale	English	German
M+	independent	völlig unabhängig
	active	sehr aktiv
	decisive	fälle leicht Entscheidungen
	never gives up	gebe nie leicht auf
	self-confident	sehr selbstsicher
	feels superior	fühle mich überlegen
	stands up under pressure	kann Druck gut standhalten
M-	aggressive	sehr aggressiv
	boastful	sehr prahlerisch
	arrogant	sehr überheblich
	greedy	sehr gierig
	dictatorial	sehr diktatorisch
	cynical	sehr zynisch
	unprincipled	halte mich an keine Grundsätze
	hostile	sehr feindselig
	egotistical	nur auf mich selbst bedacht

*Note.* M+ socially desirable masculinity. M- socially undesirable masculinity.

Table 1

*Demographic information of the sample presented separately for girls and boys*

	Girls		Boys		<i>t</i> (109)	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Age	14.34	0.54	14.58	0.65	-2.2	.030
	Girls		Boys		$\chi^2(1)$	<i>p</i>
	<i>n</i>	%	<i>n</i>	%		
Mother tongue German	39	60%	41	72%	1.63	.202
Advanced course in German	53	82%	32	56%	9.27	.002
Advanced course in math	48	74%	47	82%	1.31	.253

*Note.* *n* = 65 girls, *n* = 57 boys. One girl did not indicate her mother tongue.

Table 2

*Means (M) and standard deviations (SD) of the masculinity-scales, implicit gender-stereotyping of school, and grades in German and math presented separately for girls and boys*

	Girls		Boys		<i>t</i> (120)	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
M+	3.93	0.72	4.36	0.74	-3.20	.002	-0.58
M-	2.52	0.65	2.66	0.66	-1.12	.264	-0.20
<i>d'</i> school- female	4.13	1.12	3.58	1.13	2.67	.009	0.48
<i>d'</i> school-male	3.69	1.05	3.66	1.20	0.13	.901	0.02
<i>d'</i> -difference	0.44	1.21	-0.08	1.33	2.25	.026	0.41
Grade German	2.95	1.04	3.51	1.24	-2.66	.009	-0.49
Grade math	3.11	1.24	3.06	1.14	0.21	.84	0.04

*Note.*  $n = 65$  girls,  $n = 57$  boys. M+ socially desirable masculinity, M- socially undesirable masculinity. Both scales range from 1 (*not at all*) to 6 (*very*). *d'* indicates the strength of association between the concepts and ranges from  $-\infty$  to  $+\infty$ . Positive *d'*-difference scores indicate a stronger association of school with female than with male. Negative *d'*-difference scores indicate a stronger association of school with male than with female. Grades range from 1 (*outstanding performance*) to 7 (*insufficient performance*) with course levels (basic or advanced course) taken into account.

Table 3

*Intercorrelations between masculinity-scales, implicit gender-stereotyping of school, and grades in German and math presented separately for boys (above diagonal) and girls (below diagonal)*

	1	2	3	4	5	6	7
1 M+		.29*	-.06	.20	-.24	.12	-.04
2 M-	-.26*		-.24	-.05	-.15	.36**	.22
3 <i>d'</i> school-female	.26*	-.39**		.35**	.54**	-.06	-.04
4 <i>d'</i> school-male	.15	-.16	.38**		-.61**	-.23	-.16
5 <i>d'</i> -difference	.11	-.22	.60**	-.52**		.16	.11
6 Grade German	.06	.09	-.26*	-.07	-.18		.28*
7 Grade math	.04	-.06	.00	-.07	.06	.39**	-

*Note.*  $n = 65$  girls,  $n = 57$  boys. M+ socially desirable masculinity, M- socially undesirable masculinity, *d'* strength of association between the concepts. Positive *d'*-difference scores indicate a stronger association of school with female than with male. Negative *d'*-difference scores indicate a stronger association of school with male than with female. Grades range from 1 (*outstanding performance*) to 7 (*insufficient performance*) with course levels (basic or advanced course) taken into account.

\*  $p < .05$ , \*\*  $p < .01$

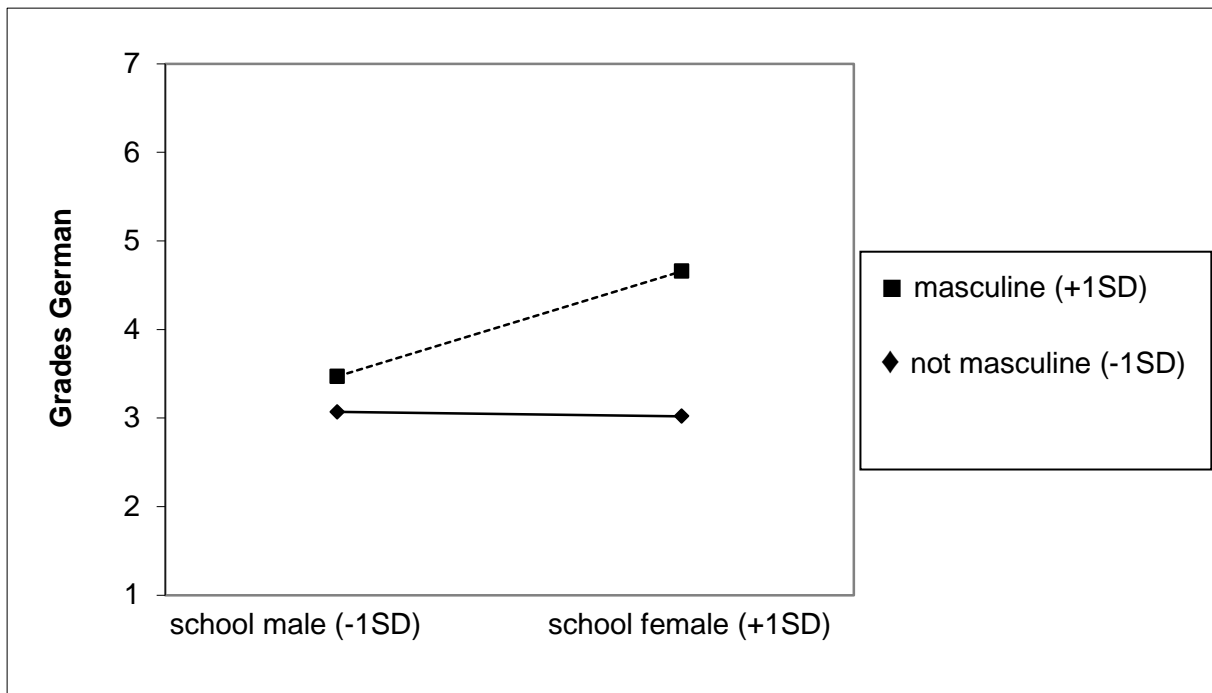
Table 4

*Hierarchical regression analyses of boys' and girls' grades in German and in math on implicit gender stereotyping of school, socially undesirable masculine traits, and the interaction between them*

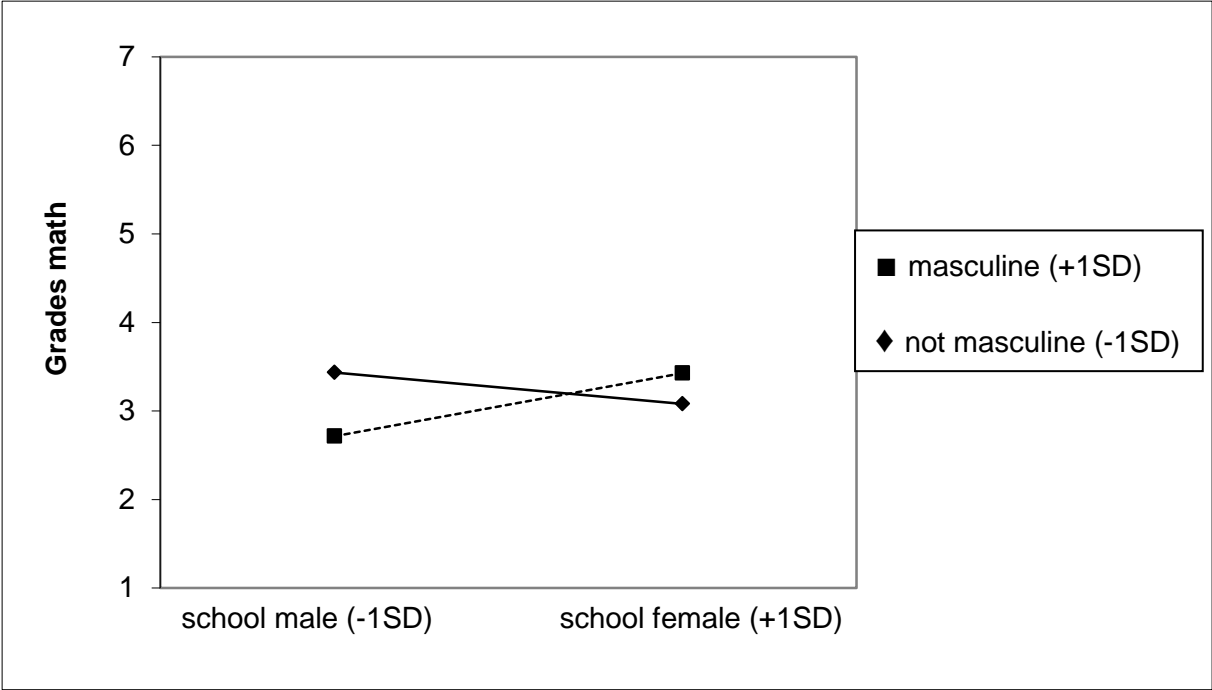
Grades in German							
Boys							
Predictors	Step 1			Step 2			
	<i>B</i>	<i>SE B</i>	$\beta$	<i>B</i>	<i>SE B</i>	$\beta$	
<i>d'</i> -difference	.20	.12	.22	.19	.11	.20	
M-	.75	.24	.40**	.87	.24	.46**	
<i>d'</i> -difference x M-				.35	.17	.25*	
$R^2 / \Delta R^2$			.18			.24 / .06	
Girls							
Predictors	Step 1			Step 2			
	<i>B</i>	<i>SE B</i>	$\beta$	<i>B</i>	<i>SE B</i>	$\beta$	
<i>d'</i> -difference	-.14	.11	-.17	-.16	.11	-.19	
M-	.09	.20	.05	.14	.21	.09	
<i>d'</i> -difference x M-				-.14	.13	-.14	
$R^2 / \Delta R^2$			.04			.05 / .02	
Grades in math							
Boys							
Predictors	Step 1			Step 2			
	<i>B</i>	<i>SE B</i>	$\beta$	<i>B</i>	<i>SE B</i>	$\beta$	
<i>d'</i> -difference	.12	.11	.14	.13	.11	.15	
M-	.42	.23	.24	.35	.24	.20	
<i>d'</i> -difference x M-				-.18	.18	-.14	
$R^2 / \Delta R^2$			.07			.09 / .02	
Girls							
Predictors	Step 1			Step 2			
	<i>B</i>	<i>SE B</i>	$\beta$	<i>B</i>	<i>SE B</i>	$\beta$	
<i>d'</i> -difference	.05	.13	.05	.10	.13	.09	
M-	-.09	.25	-.05	-.23	.25	-.12	
<i>d'</i> -difference x M-				.34	.16	.28*	
$R^2 / \Delta R^2$			.01			.08 / .07	

*Note.*  $n = 65$  girls,  $n = 57$  boys. M- socially undesirable masculinity, *d*-difference gender stereotyping of school. Positive *d'*-difference scores indicate a stronger association of school with female than with male. Negative *d'*-difference scores indicate a stronger association of school with male than with female. Grades range from 1 (*outstanding performance*) to 7 (*insufficient performance*) with course levels (basic or advanced course) taken into account.

\*  $p < .05$ , \*\*  $p < .01$



**Fig. 1** Boys' grades in German as a function of socially undesirable masculinity and gender stereotyping of school. Grades ranging from 1 (*outstanding performance*) to 7 (*insufficient performance*) with course levels (basic or advanced course) taken into account



**Fig. 2** Girls' grades in math as a function of socially undesirable masculinity and gender stereotyping of school. Grades ranging from 1 (*outstanding performance*) to 7 (*insufficient performance*) with course levels (basic or advanced course) taken into account





### **5.3 Studie 3: Do teachers equate male and masculine with lower academic engagement? How student's gender enactment triggers gender stereotypes at school**

Heyder, A. & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How student's gender enactment triggers gender stereotypes at school. *Social Psychology of Education*. Advance online publication. doi: 10.1007/s11218-015-9303-0

The final publication is available at Springer via <http://dx.doi.org/10.1007/s11218-015-9303-0>.

The current research was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) allocated to the second author (KE 1412/2-1).

### Abstract

Girls presently outperform boys in overall academic success. Corresponding gender stereotypes portray male students as lazy and troublesome and female students as diligent and compliant. The present study investigated whether these stereotypes impact teachers' perceptions of students and whether students' visible enactment of their gender at school (behaving in a very masculine or feminine way) increases the impact of these stereotypes on teachers' perceptions of students. We hypothesized that teachers would ascribe more behavior that impedes learning and less behavior that fosters learning to male students who enact masculinity as compared with male students who show gender-neutral behavior and female students. Three pilot studies ( $N = 104$ ;  $N = 82$ ;  $N = 86$ ) yielded pretested material for a randomized vignette study of  $N = 104$  teachers. The teachers read one randomly assigned vignette describing a male (or female) student enacting his (or her) gender (or not) and rated how likely this student would be to display behaviors that impede or foster learning in a 2 (between: target students' gender) x 2 (between: gender enactment [yes/no]) x 2 (between: teachers' gender) x 2 (within: ascribed behavior) factorial design. As expected, male students enacting masculinity were rated as showing the lowest amount of academic engagement. Results are discussed with regard to the current debate on the causes of boys' lower academic success.

Keywords: gender stereotypes, masculinity, gender differences in academic achievement, teachers' expectations, academic engagement

## 6 Gesamtdiskussion

Die vorliegende Arbeit hat das Ziel, zur Erklärung des differenziellen Schulerfolgs von Jungen und Mädchen beizutragen. Auf Grundlage des bisherigen Forschungsstands zur Erklärung von Geschlechtsunterschieden im Schulerfolg wurden folgende Fragestellungen abgeleitet und in drei Studien empirisch geprüft:

- (1) Können bedeutende Variablen des Eccles et al.-Modells, d.h. FSK, Werte, Elterneinschätzung der kindlichen Fähigkeiten, Intelligenz, Bildungsstand der Eltern und die früheren Leistungen, die Geschlechtsunterschiede nicht nur in den Mathematik- sondern auch in den Deutschnoten erklären?
- (2) Ist Schule implizit feminin stereotypisiert und welche Rolle spielt dies für den differenziellen Schulerfolg von Jungen und Mädchen?
- (3) Kann Gender Enactment von Jugendlichen bei Lehrkräften stereotypisierte Erwartungen zu Ungunsten des Lernverhaltens und des Benehmens von Jungen auslösen?

Im Folgenden werden zunächst die Ergebnisse der einzelnen Studien hinsichtlich der aufgestellten Fragestellung zusammengefasst. Anschließend werden studienübergreifende Aspekte der Ergebnisse sowie ihre Limitationen diskutiert und dargelegt, welche Implikationen sich aus den Ergebnissen für die weitere Forschung und die schulische Praxis ableiten lassen. Die Arbeit endet mit einem studienübergreifenden Fazit.

### 6.1 Zusammenfassung der Studienergebnisse

Studie 1 (*Beyond expectancies and values: Applying the Eccles et al. model to the underachievement of boys*) verfolgte das Ziel, die Geschlechtsunterschiede in den Mathematiknoten zu Gunsten der Jungen und in den Deutschnoten zu Gunsten der Mädchen mithilfe des Eccles al.-Modells (Eccles et al., 1983; Wigfield & Eccles, 2000) zu erklären. Zu diesem Zweck wurden schulerfolgsrelevante individuelle Merkmale der Schülerin oder des

Schülers (Werte für Mathematik und Deutsch, FSK in Bezug auf Mathematik und Deutsch, frühere Mathematik- und Deutschnote, Intelligenz) und schulerfolgsrelevante Merkmale seiner oder ihrer sozialisierenden Umwelt (Bildungsstand der Eltern, Elterneinschätzung der Fähigkeiten des Kindes in Mathematik und Deutsch) gemäß den theoretischen Annahmen des Modells als mögliche Mediatoren des Effekts von Geschlecht auf die Note in zwei fachspezifische Strukturgleichungsmodelle aufgenommen. In Mathematik konnten die genannten Variablen den Geschlechtsunterschied in den Noten nicht nur erklären, sondern übererklären: Wurden die genannten Merkmale als vermittelnde Variablen berücksichtigt, kehrte sich der Geschlechtsunterschied um. Denn die höhere Wertschätzung des Fachs Mathematik durch die Jungen, ihre besseren früheren Mathematiknoten, höhere Intelligenz und positivere Einschätzung ihrer mathematischen Fähigkeiten durch die Eltern wirkten als Suppressorvariablen gemäß des Verständnisses von MacKinnon, Krull und Lockwood (2000). Die besseren Noten der Mädchen im Fach Deutsch hingegen konnten nicht erklärt werden. Zwar wurde der indirekte Effekt über die besseren Noten der Mädchen zum früheren Messzeitpunkt statistisch signifikant, er wurde jedoch durch den indirekten Effekt über die in dieser Gymnasialstichprobe höhere Intelligenz der Jungen ausgeglichen. Insgesamt zeigten die Mediationsmodelle in *beiden* untersuchten Fächern direkte Effekte von Geschlecht auf Noten – und damit schlechtere Noten der Jungen – in vergleichbarer Größe auf, welche durch die untersuchten Variablen des Eccles et al.-Modells nicht erklärt werden konnten.

Wegen des geringeren schulischen Lernverhaltens und Engagements der Jungen im Vergleich zu Mädchen und dessen Bedeutung für die Erklärung von Geschlechtsunterschieden im Schulerfolg untersuchte Studie 2 (*Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement*) mithilfe des GNAT (Nosek & Banaji, 2001), ob Schule und schulisches Lernen implizit als feminin wahrgenommen werden, und wie dies mit dem Schulerfolg von Jungen und Mädchen zusammenhängt. Dabei zeigte sich entgegen manch medialer Darstellung, dass Schule und

schulisches Lernen im Durchschnitt nur leicht feminin wahrgenommen wurden. Die Effektgröße war klein. Jungen assoziierten Schule und Lernen im Durchschnitt implizit gleichstark mit männlich und mit weiblich, Mädchen hingegen deutlich stärker mit ihrem eigenen Geschlecht. Damit zeigte Studie 2 auf, dass Schule insgesamt deutlich weniger eindeutig geschlechtsstereotypisiert wahrgenommen wird als einzelne Schulfächer (siehe z.B. Nosek & Smyth, 2011). Gemäß den theoretischen Annahmen des IIRM (Kessels & Hannover, 2004, 2006) und der IBMT (Elmore & Oyserman, 2012; Oyserman & Destin, 2010) war nicht die Wahrnehmung von Schule und Lernen als geschlechtsstereotypisiert allein prädiktiv für Schulerfolg, sondern die aus der Stereotypisierung resultierende Passung zur eigenen geschlechtsbezogenen Identität. Denn eine Schule-weiblich-Assoziation ging nur bei den Jungen, die sich selbst stark mit sozial unerwünschten maskulinen Eigenschaften beschrieben, mit schlechteren Noten im ebenfalls feminin stereotypisierten Fach Deutsch einher. Mit den Mathematiknoten der Jungen oder den Mathematik- oder Deutschnoten der Mädchen zeigten sich keine signifikanten Zusammenhänge. Diese Ergebnisse betonen die Bedeutung der wahrgenommenen Passung zwischen der Identität von Jugendlichen und schulischen Lernangeboten und deuten auf ein komplexes Zusammenspiel zwischen geschlechtsbezogenem Selbstkonzept, Geschlecht, fachlichen und verhaltensbezogenen Geschlechterstereotypen und Schulerfolg hin.

Aus früheren Studien mit Lehrkräften ist bekannt, dass die Anstrengung und das Lernverhalten, das Lehrkräfte ihren Schülerinnen und Schülern zuschreiben, sich positiv auf die vergebenen Noten auswirken (z.B. McMillan, 2001). Aufgrund dieser Relevanz des Lernverhaltens für die Notengebung untersuchte die dritte Studie (*Do teachers equate male and masculine with lower academic engagement? How student's gender enactment triggers gender stereotypes at school*) im Rahmen eines Vignettenexperiments, ob sich Gender Enactment von Schülerinnen und Schülern auf das ihnen von Lehrkräften zugeschriebene Lernverhalten auswirkt. Hypothesenkonform erwarteten Lehrkräfte grundsätzlich weniger

lernförderliches und mehr lernhinderliches Verhalten von Jungen als von Mädchen, auch wenn beide identisches, geschlechtsneutrales Verhalten zeigten und sich nur in ihren Vornamen unterschieden. Der erste, spontane Eindruck der Lehrkräfte spiegelte folglich das Stereotyp des faulen Jungen und das des fleißigen Mädchens wider. Entsprechend der aus dem Kontinuum-Modell der Eindrucksbildung (Fiske & Neuberg, 1990) abgeleiteten Hypothesen verstärkte Gender Enactment den Einfluss dieser verhaltensbezogenen Geschlechterstereotype auf die Eindrucksbildung von Lehrkräften zu Ungunsten der Jungen. Lehrkräfte schrieben einem Jungen, der seine Männlichkeit inszenierte, ein schlechteres Lernverhalten zu, als einem Jungen, der sich neutral verhielt. Jedoch erwarteten Lehrkräfte auch von Mädchen, die ihre Femininität inszenierten, ein schlechteres Lernverhalten als von sich neutral verhaltenden Mädchen. Das Mädchen, das seine Femininität inszenierte, wurde aber hinsichtlich seines Lernverhaltens von den Lehrkräften immer noch positiver eingeschätzt als der Junge, der seine Maskulinität inszenierte.

## **6.2 Studienübergreifende Aspekte**

Die Ergebnisse der drei Studien weisen auf mehrere gemeinsame Aspekte hin, die bei der Erklärung von Geschlechtsunterschieden im Schulerfolg bedeutsam zu sein scheinen: (1) Geschlechterstereotype, (2) die geschlechtsbezogene Identität und Identitätsentwicklung Jugendlicher, (3) das betrachtete Fach und (4) der Schulerfolgsindikator Noten. Diese studienübergreifenden Aspekte werden in den folgenden Abschnitten vor dem Hintergrund weiterer Forschungsergebnisse interpretiert und diskutiert.

### **6.2.1 Bedeutung von Geschlechterstereotypen**

Geschlechterstereotype können – wie alle Stereotype – die Informationsverarbeitung und das Verhalten von Menschen in sozialen Interaktionen beeinflussen (z.B. Fiedler & Bless, 2003), auch im schulischen Kontext. Sie können sowohl auf die Person wirken, die der Gruppe

angehört, auf die sich das Stereotyp bezieht (z.B. Nguyen & Ryan, 2008; Spencer, Steele & Quinn, 1999), als auch auf die Personen, die sie wahrnehmen und mit denen sie interagiert (z.B. Fiske & Neuberg, 1990). Im Rahmen der vorliegenden Arbeit wurde die Bedeutung einer stereotypisierten Wahrnehmung von Schule und schulischem Lernen als feminin (1) für den Schulerfolg von Schülerinnen und Schülern und (2) für die Eindrucksbildung von Lehrkräften untersucht. Im Kontext der Pädagogischen Psychologie hatten bisher vor allem die fachspezifischen Geschlechterstereotype große Aufmerksamkeit als Einflussfaktoren auf Leistung (z.B. Nguyen & Ryan, 2008; Spencer et al., 1999) und Motivation (z.B. Eccles et al., 1983; Kessels & Hannover, 2004; Oyserman & Destin, 2010; Plante, de la Sablonnière, Aronson & Théorêt, 2013) erfahren. Die Studien 2 und 3 der vorliegenden Arbeit untersuchten hingegen die Bedeutung von verhaltensbezogenen und damit fachunspezifischen Geschlechterstereotypen und griffen damit die öffentliche Wahrnehmung und mediale Darstellung von Jungen als schulischen Problemfällen und Bildungsverlierern (Hartley & Sutton, 2013; Latsch & Hannover, 2014) auf.

Mit Studie 3 gelang es im Rahmen eines Vignettenexperimentes nachzuweisen, dass auch Jugendliche durch Verhaltensweisen, welche die eigene Geschlechtszugehörigkeit betonen, verhaltensbezogene Geschlechterstereotype aktivieren und ihren Effekt auf die Eindrucksbildung bei Lehrkräften verstärken können. Frühere Studien hatten vielfach aufgezeigt, dass situative Merkmale oder das Verhalten der Lehrkraft selbst die Salienz der Kategorie Geschlecht erhöhen und damit Geschlechterstereotype und stereotypenkonformes Verhalten bei Schülerinnen und Schülern auslösen können (z.B. Bigler, 1995; Hilliard & Liben, 2010; Shih, Pittinsky & Ambady, 1999; Steele & Aronson, 1995). Studie 3 wies nach, dass auch die andere Wirkrichtung funktioniert: Das Verhalten der Schülerinnen und Schüler löste stereotypenbasierte Erwartungen bei Lehrkräften aus. Da verhaltensbezogene Stereotype Jungen aktuell als faul und wenig angepasst und Mädchen als fleißig und angepasst beschreiben, wirkte dieser Mechanismus besonders zum Nachteil der Jungen, die ihre

Männlichkeit in der Schule betonten. Ihnen schrieben die Lehrkräfte am wenigsten lernförderliches und am meisten lernhinderliches Lernverhalten zu. Solch ein Eindruck kann sich mittelbar in schlechteren Noten und Leistungsbeurteilungen niederschlagen, wie die umfangreiche Literatur zur Notengebungspraxis (z.B. Brookhart, 1993; McMillan, 2001; Randall & Engelhard, 2010) und zur Bedeutung von Anstrengung und angepasstem Verhalten für das Erreichen guter Noten und positiver Leistungseinschätzungen (z.B. DiPrete & Jennings, 2012; Downey & Vogt Yuan, 2005; Kaiser, Retelsdorf, Südkamp & Möller, 2013; McDermott, Mordell & Stoltzfus, 2001) belegt.

Die Bedeutung von Geschlechterstereotypen für den differenziellen Schulerfolg von Jungen und Mädchen wird durch die Ergebnisse von Studie 2 in einer Schülerstichprobe unterstrichen. Denn je stärker maskuline Jungen Schule und schulisches Lernen implizit als feminin stereotypisiert wahrnahmen, desto schlechtere Noten erzielten sie im ebenfalls feminin stereotypisierten Schulfach Deutsch. Dieses Ergebnis zeigt, dass die Geschlechtsstereotypisierung von Schule und schulischem Lernen mit der Geschlechtsstereotypisierung von Schulfächern sowie mit dem geschlechtsbezogenen Selbstkonzept der Jugendlichen interagieren kann. Fachliche und überfachliche Geschlechterstereotype sind nicht nur, wie vielfach nachgewiesen wurde, einzeln bedeutend für die Erklärung von Geschlechtsunterschieden im Schulerfolg (z.B. Elmore & Oyserman, 2012; Nosek, Banaji & Greenwald, 2002; Steffens, Jelenec & Noack, 2010), sondern entfalten ihre volle Bedeutung vor allem im Zusammenspiel. Die Bedeutung der Interaktion fachunspezifischer und fachspezifischer Geschlechterstereotype wiesen auch Latsch und Hannover (2014) nach. In ihren Studien ging eine Bedrohung durch das fachunspezifische Jungen-als-Bildungsverlierer-Stereotyp bei Jungen – ähnlich wie in Studie 2 dieser Arbeit – nur mit Leistungseinbußen in der feminin stereotypisierten Domäne Lesen einher, nicht jedoch in der maskulin stereotypisierten Domäne Mathematik.



Die robusten, vielfach nachgewiesenen, fachspezifischen Geschlechterstereotype könnten potentiell auch eine inhaltliche Erklärung der Ergebnisse von Studie 1 in Mathematik bieten. Denn die stereotype Annahme, dass Jungen eine höhere mathematische Begabung und Leistungsfähigkeit besäßen als Mädchen (z.B. Steffens et al., 2010), könnte dazu geführt haben, dass die Jungen die eigenen mathematischen Fähigkeiten und Eltern die mathematischen Fähigkeiten ihres Sohnes im Vergleich zu seinen tatsächlichen Leistungen in Studie 1 überschätzt haben. Bekanntlich haben verschiedene Studien bereits nachgewiesen, dass das Mathematik-männlich-Stereotyp positiv mit dem mathematischen FSK von Jungen korreliert (z.B. Steffens et al., 2010) und dass Jungen die eigenen mathematischen Fähigkeiten und Eltern die mathematischen Fähigkeiten ihres Sohnes überschätzen (z.B. Frome & Eccles, 1998; Herbert & Stipek, 2005). Tatsächlich fielen in Studie 1 die Geschlechtsunterschiede im mathematischen FSK und in den Elterneinschätzungen deutlich größer aus, als die in den tatsächlich vergebenen Mathematiknoten. Solche potentiellen Überschätzungen der mathematischen Fähigkeiten der Jungen könnten wiederum miterklären, warum Jungen unter Kontrolle der genannten Variablen in Studie 1 schlechtere Mathematiknoten erhielten als Mädchen. Eine empirische Prüfung dieser Erklärung steht jedoch noch aus.

Zusammenfassend zeigen die Ergebnisse von Studie 2 und 3 auf, dass die stereotype Verknüpfung von Schule und schulischem Lernverhalten mit „Mädchen“ und „feminin“ eine weitere wichtige Erklärung für den geringeren Schulerfolg mancher Jungen darstellt. Die Bedeutung dieser Geschlechtsstereotypisierung wurde im Rahmen der vorliegenden Arbeit sowohl in einer Stichprobe bestehend aus Schülerinnen und Schüler als auch in einer Stichprobe bestehend aus Lehrkräften aufgezeigt und kann, wie in Studie 2 nachgewiesen, mit fachspezifischen Geschlechterstereotypen interagieren. Fachspezifische Geschlechterstereotype könnten auch eine mögliche inhaltliche Erklärung für die Umkehrung

des Geschlechtsunterschieds in den Mathematiknoten im Mediationsmodell in Studie 1 darstellen.

### **6.2.2 Bedeutung der geschlechtsbezogenen Identität**

Die Entwicklung und Ausdifferenzierung der eigenen geschlechtsbezogenen Identität ist eine zentrale Entwicklungsaufgabe im Jugendalter (Finkenauer, Engels, Meeus & Oosterwegel, 2002; siehe auch Erikson, 1968; Havighurst, 1956), die mit Schulerfolg im Konflikt stehen kann (Hannover & Kessels, 2011; Kessels, Heyder, Latsch & Hannover, 2014). Studie 2 und 3 wiesen am Beispiel der Passung zwischen Identität und schulischen Angeboten und am Beispiel des Gender Enactments die Bedeutung der geschlechtsbezogenen Identität für die Erklärung des differenziellen Schulerfolgs von Jungen und Mädchen nach. Wie auf Grundlage des IIRM und der IBMT erwartet konnte das Ausmaß der Passung zwischen der geschlechtsbezogenen Identität der Jugendlichen und Schule und schulischem Lernen in Studie 2 den Schulerfolg von Jungen vorhersagen: Je femininer sie Schule wahrnahmen und je stärker sie sich sozial unerwünschte maskuline Eigenschaften zuschrieben, desto schlechtere Noten erhielten sie im ebenfalls feminin stereotypisierten Fach Deutsch. Damit wurde in Studie 2 auch empirisch nachgewiesen, dass die Passung zwischen Identität und schulischen Angeboten nicht nur einen Beitrag zur Erklärung der Unterrepräsentanz der Mädchen im MINT-Bereich leisten kann, sondern auch zur Erklärung des geringeren Schulerfolgs der Jungen.

Studie 3 untersuchte die Wirkung von Gender Enactment auf die Eindrucksbildung von Lehrkräften. Aus Studien zur Stereotypisierung von Schulfächern ist bekannt, dass die direkte Distanzierung von schulischen Angeboten, die als nicht passend zur eigenen geschlechtsbezogenen Identität wahrgenommen werden, eine Möglichkeit des Gender Enactments darstellt (z.B. Kessels, Warner, Holle & Hannover, 2008). Doch Gender Enactment kann auch durch Verhaltensweisen geschehen, die keinen direkten Bezug zu

Schule und Lernen aufweisen (Eckes & Trautner, 2000; Hilliard & Liben, 2010). Im Rahmen von Studie 3 wurde in einem experimentellen Design nach Kenntnis der Autorin erstmalig nachgewiesen, dass auch die Inszenierung der eigenen Geschlechtsidentität mit schul- und lernunabhängigen Verhaltensweisen stereotypenbasierte, negative Konsequenzen für den Schulerfolg von Mädchen und Jungen – aber besonders für den Schulerfolg von Jungen – auslösen kann. Dieser Befund veranschaulicht, dass die Identitätsentwicklung von Jugendlichen nicht nur über die Bevorzugung identitätskongruenter schulischer Angebote (z.B. Kessels, Rau & Hannover, 2006; Steffens et al., 2010), sondern auch durch Gender Enactment und Stereotypenaktivierung bei Lehrkräften zum differenziellen Schulerfolg von Jungen und Mädchen beitragen kann.

Außerdem weisen sowohl die Ergebnisse von Studie 2 als auch von Studie 3 darauf hin, dass es nicht reicht, nur das biologische Geschlecht der Jugendlichen an sich heranzuziehen, um die Bedeutung der geschlechtsbezogenen Identität für die Erklärung von Unterschieden im Schulerfolg zu ermitteln, sondern dass innerhalb der Gruppe der Mädchen und der Gruppe der Jungen stärker differenziert werden muss (siehe auch Kessels & Steinmayr, 2013b; Lyng, 2009). Denn für Jungen mit einem sehr maskulinen geschlechtsbezogenen Selbstkonzept (Studie 2) und für Jungen, die ihre männliche Identität betonten (Studie 3), zeigte sich ein ungünstigeres Bild hinsichtlich ihrer Noten und ihres zugeschriebenen Lernverhaltens als für weniger maskuline oder sich neutral verhaltende Jungen. Die Ergebnisse von Studie 3 zeigen darüber hinaus auf, dass auch Gender Enactment von Mädchen negative Erwartungen bezüglich ihres Lernverhaltens auslösen kann. Eine wahrscheinliche Erklärung für dieses zunächst unerwartete Ergebnis ist, dass aufgrund der Kontrolle des Stimulusmaterials für etwaige Unterschiede in der sozialen Erwünschtheit (siehe S. 186-188). Verhaltensweisen zur Inszenierung von Femininität ausgewählt wurden, die ein *Substereotyp* über eine spezifische Gruppe von Mädchen aktiviert haben, und zwar das der sehr femininen aber auch schuldistanten „Babes“. Als „Babes“ war in einer ethnografischen Studie zu verschiedenen

Schülertypen die Gruppe der Mädchen bezeichnet worden, die sich einerseits durch Kleidung, Make-Up etc. sehr feminin gaben und andererseits schulisches Engagement ablehnten (Lyng, 2009). Eine Aktivierung des Babes-Stereotyps könnte folglich zu den negativeren Erwartungen bezüglich des Lernverhaltens führen, die sich in Studie 3 gezeigt haben. Gestärkt wird diese Vermutung ebenfalls durch den Zusammenhang zwischen einer gesteigerten Fokussierung auf die eigene Weiblichkeit und auf das eigene Aussehen und einem geringeren Schulerfolg bei Mädchen, den McKenney und Bigler (2014) mit „high heels, low grades“ auf den Punkt brachten.

Insgesamt lässt sich folglich festhalten, dass je nachdem, wie die geschlechtsbezogene Identität der Jugendlichen ausgestaltet und konstruiert wird, sie *sowohl für Mädchen als auch für Jungen* unterschiedlich gut zu schulischem Lernen und Schulerfolg passt. Dabei kann sie nicht nur Geschlechtsunterschiede in den fachspezifischen Interessen und Leistungen erklären, sondern, wie in Studie 2 und 3 gezeigt, auch einen wichtigen Beitrag zur Erklärung des geringeren Schulerfolgs der Jungen insgesamt leisten.

### **6.2.3 Bedeutung des untersuchten Fachs**

Sowohl in Studie 1 als auch in Studie 2 unterschieden sich die Ergebnisse in Abhängigkeit davon, welches Fach betrachtet wurde. Die untersuchten Variablen des Eccles et al.-Modells konnten die schlechteren Noten der Jungen in Deutsch nicht erklären und dafür ihre besseren Mathematiknoten übererklären (Studie 1). Die Interaktion zwischen einer Schule-weiblich-Assoziation und einem maskulinen Selbstkonzept ging bei Jungen mit schlechteren Noten im Fach Deutsch, nicht aber im Fach Mathematik, einher (Studie 2). Diese unterschiedlichen Ergebnisse für beide Domänen stellen klar, dass die Ergebnisse in einer Domäne nicht ohne weiteres auf eine andere Domäne übertragen werden können. Außerdem werfen sie die Frage auf, auf welche (noch nicht untersuchten) fachspezifischen Faktoren diese differenziellen Effekte zurückgeführt werden könnten. Eine in Abschnitt 6.2.1 bereits dargelegte Erklärung

bieten die fachspezifischen Geschlechterstereotype. Alternativ könnten die differenziellen Effekte für Mathematik und Deutsch in Studie 1 und 2 möglicherweise auch zu einem Teil darauf zurückzuführen sein, dass Noten in beiden Fächern nicht genau das Gleiche abbilden. Leider haben bisher nur sehr wenige der vielen Studien zur Notengebungspraxis zwischen Lehrkräften verschiedener Fächer differenziert (McMillan, 2001; McMillan, Myran & Workman, 2002; Resh, 2009). Resh (2009) befragte die Lehrkräfte der israelischen PISA-Stichprobe gezielt nach dem relativen Gewicht, das sie den Faktoren Leistung, Anstrengung und Unterstützungsbedürfnis einer Schülerin oder eines Schülers bei der Notengebung beimessen, und deckte interessante Fachunterschiede auf: Die Mathematiklehrkräfte berücksichtigten stärker die gezeigte Leistung der Schülerinnen und Schüler und weniger stark die Anstrengung als diejenigen Lehrkräfte, die die Muttersprache oder Naturwissenschaften unterrichteten. Zu ähnlichen Ergebnissen kam auch eine US-amerikanische Studie (McMillan, 2001). Dies könnte mit fachspezifischen Annahmen darüber, welche Eigenschaften und Verhaltensweisen für den Erfolg in einem Fach wichtig sind, zusammenhängen. In manchen universitären Fachkulturen, z.B. Philosophie und Mathematik, gilt stärker die Annahme, dass Erfolg in dem Fach eine natürliche, „angeborene“ Begabung voraussetze und Anstrengung „allein“ nicht reiche um Spitzenleistungen zu bringen, als in anderen Fächern, z.B. Psychologie (Leslie, Cimpian, Meyer & Freeland, 2015). Ob sich jedoch solche oder andere fachspezifische Faktoren in Unterschieden zwischen Fächern darin äußern, was Noten abbilden und wodurch interindividuelle Unterschiede in ihnen zu erklären sind, bleibt eine spannende Frage für die zukünftige Forschung. Die differenziellen Effekte für Mathematik und Deutsch in Studie 1 und 2 betonen jedoch, dass es wichtig ist, zwischen beiden Domänen zu differenzieren.

#### **6.2.4 Bedeutung des Schulerfolgsindikators Noten**

Da Noten für Schülerinnen und Schüler die bedeutendste Leistungsrückmeldung darstellen und über ihren weiteren Bildungsverlauf unmittelbar mitentscheiden (z.B. Ziegenspeck, 1999), wurden sie in Studie 1 und 2 als Indikator für Schulerfolg herangezogen. Während sich im Fach Deutsch sowohl in Studie 1 als auch in Studie 2 Notenunterschiede zu Gunsten der Mädchen zeigten, berichteten die Jungen in Studie 1 bessere und in Studie 2 gleich gute Mathematiknoten wie die Mädchen. Dieses konsistente Bild in der Muttersprache und etwas uneinheitlichere Bild in Mathematik in Deutschland entspricht bisherigen Studien aus dem Sekundarschulbereich (z.B. Dumont et al., 2012; Steinmayr & Spinath, 2008; Voyer & Voyer, 2014; siehe Kapitel 1.2).

Auf rein deskriptiver Ebene fällt außerdem auf, dass die Geschlechtsunterschiede in den Schulnoten sich zwischen der Gesamtschulstichprobe aus Studie 1 und der Gymnasialstichprobe aus Studie 2 unterscheiden. Die Geschlechtsunterschiede in der Deutschnote zu Gunsten der Mädchen fielen in der Gesamtschulstichprobe etwas größer aus als in der Gymnasialstichprobe. In Mathematik erreichten die Jungen in der Gymnasialstichprobe hingegen bessere Noten, in der Gesamtschulstichprobe nicht. Dieses Muster könnte darauf hindeuten, dass sich die Abhängigkeit der Größe der gefundenen Geschlechtsunterschiede in Kompetenzen von der untersuchten Schulform (Hosenfeld, Köller & Baumert, 1999; Stanat & Kunter, 2001; siehe Kapitel 1.3) tendenziell auch in Geschlechtsunterschieden in den Noten widerspiegelt: Werden nur die Schülerinnen und Schüler betrachtet, die ein Gymnasium besuchen, zeigen sich größere Geschlechtsunterschiede in Mathematiknoten und -kompetenzen zu Gunsten der Jungen und kleinere Geschlechtsunterschiede in den Noten und Kompetenzen der verbalen Domäne zu Gunsten der Mädchen, als wenn die Gesamtstichprobe untersucht wird. Da in Studie 1 und 2 jedoch nicht nur zwei Gelegenheitsstichproben, sondern auch verschiedene Jahrgangsstufen

untersucht wurden, ist diese ausschließlich deskriptive Beobachtung mit großer Vorsicht zu betrachten.

Die Ergebnisse von Studie 1 weisen außerdem darauf hin, dass sich hinter auf den ersten Blick besseren Mathematiknoten der Jungen auch ein „relatives Underachievement“ ihrerseits im Sinne eines Nichtausschöpfens ihres Potentials verbergen kann. Das multiple Mediationsmodell zur Erklärung der besseren Mathematiknoten der Jungen deckte auf, dass sich unter Kontrolle der individuellen Merkmale Intelligenz, FSK, Werte und frühere Leistung sowie der Elterneinschätzungen der kindlichen Fähigkeiten und des Bildungsstands der Eltern der Geschlechtseffekt umkehrt. Unterschieden sich ein Mädchen und ein Junge auf den genannten Variablen nicht, erhielt das Mädchen auch in Mathematik bessere Noten als der Junge. Während bisherige Studien vor allem auf die Diskrepanz zwischen dem Abschneiden von Jungen und Männern in standardisierten Leistungs- oder Intelligenztests und ihren erreichten Noten hinwiesen (z.B. Duckworth & Seligman, 2006; Helbig, 2010; Kenney-Benson, Pomerantz, Ryan & Patrick, 2006; siehe Kapitel 1.3 und 2.1), zeigte Studie 1, dass es Jungen auch weniger gut gelingt, ihren Vorteil in den untersuchten motivationalen und sozialen Ressourcen gegenüber den Mädchen in gute Mathematiknoten zu überführen. Gleichzeitig deutet sich hier bereits an, dass bei der Analyse von Geschlechtsunterschieden im Schulerfolg Noten als alleiniger Indikator für Schulerfolg weniger aussagekräftig sind, als wenn zusätzlich das kognitive, motivationale oder soziale Potential der Schülerinnen und Schüler berücksichtigt wird. Diese Einschränkung wird im folgenden Kapitel 6.3 Limitationen näher ausgeführt.

### **6.3 Limitationen**

Die Belastbarkeit der berichteten und diskutierten Ergebnisse wird durch verschiedene methodische Merkmale der Studien eingeschränkt. Nachdem die spezifischen Limitationen der einzelnen Studien bereits ausführlich in den Diskussionsteilen der Manuskripte erörtert

wurden (Studie 1: S. 126; Studie 2: S. 161; Studie 3: S. 196), werden an dieser Stelle noch einmal studienübergreifende Limitationen aufgegriffen und kritisch diskutiert. Diese betreffen vor allem die verwendeten Stichproben und Forschungsdesigns.

Eine Herausforderung für *alle* Studien, die sich mit der Erklärung von Geschlechtsunterschieden im Schulerfolg beschäftigen, stellt grundsätzlich die Abhängigkeit der verschiedenen Indikatoren für Schulerfolg dar (siehe Kapitel 1). So sind die erworbenen Kompetenzen von Schülerinnen und Schülern wichtiger Bestandteil ihrer erreichten Noten (z.B. Duckworth, Quinn & Tsukayama, 2012; Hülür, Wilhelm & Robitzsch, 2011; Randall & Engelhard, 2010). Noten wiederum tragen ihrerseits zu Unterschieden in der Bildungsbeteiligung (z.B. Lehmann, Peek & Gänsfuß, 1997; Maaz & Nagy, 2010) und dem Erwerb von Schulabschlüssen bei (z.B. Verordnung über die Schularten und Bildungsgänge der Sekundarstufe I, 2010). Gleichzeitig fördert der Besuch einer höheren Schulform den Kompetenzerwerb stärker als der einer niedrigeren Schulform (z.B. Baumert, Stanat & Watermann, 2006). Da sich wie in Kapitel 1 dargestellt zusätzlich die Disparitäten im Schulerfolg zwischen Jungen und Mädchen unterscheiden, je nachdem welcher Indikator für Schulerfolg betrachtet wird, erfassen Studien, die nur einen Indikator berücksichtigen, nicht die volle Komplexität des Phänomens. Um angesichts dieser Schwierigkeit dennoch eine möglichst heterogene und unselektierte Stichprobe nutzen zu können, wurde Studie 2 an einer nordrhein-westfälischen Gesamtschule durchgeführt. Gesamtschulen bieten alle Schulabschlüsse bis zum Abitur an und werden in Nordrhein-Westfalen insgesamt von ungefähr gleich vielen Mädchen wie Jungen besucht (Statistisches Bundesamt, 2013). Auf diese Weise wurde versucht, die Konfundierung der Ergebnisse durch Unterschiede in der Bildungsbeteiligung zwischen Jungen und Mädchen gering zu halten. Studie 1 hingegen untersuchte gezielt Geschlechtsunterschiede in Schulnoten am Gymnasium, da Unterschiede zu Ungunsten der Jungen hier angesichts der unterschiedlichen Eingangsselektivität für Jungen und Mädchen besonders interessant und erklärungsbedürftig sind. In der Tat schnitten



die Jungen in der Gymnasialstichprobe aus Studie 1 im Intelligenztest besser ab als die Mädchen und ihre Eltern wiesen einen tendenziell (jedoch nicht signifikant) höheren Bildungsgrad auf als die Eltern der Mädchen. Dies untermauert die Annahme, dass es sich bei Jungen am Gymnasium um eine stärker selektierte Gruppe handelt als bei den Mädchen, und bestärkt im Nachhinein die Entscheidung, diese Merkmale erhoben und in den statistischen Modellen in Studie 1 berücksichtigt zu haben. Noch besser könnte den beschriebenen Abhängigkeiten der verschiedenen Indikatoren für Schulerfolg untereinander in der Forschung jedoch mit Datensätzen begegnet werden, die die drei genannten Indikatoren gleichzeitig enthalten und möglichst repräsentativ sind. Dies ist praktisch aber nur im Rahmen sehr großer Forschungsverbünde oder durch Sekundäranalysen bereits existierender repräsentativer Stichproben, wie sie beispielsweise das Nationale Bildungspanel bereitstellt, möglich. Das wiederum schränkt die untersuchbaren Fragestellungen sehr ein, da logischerweise nur die Variablen genutzt werden können, die bereits im Datensatz enthalten sind. Beispielsweise erscheint es schwer vorstellbar, dass im Rahmen einer großen Schulleistungsuntersuchung implizite Assoziationsmaße eingesetzt würden, wie sie in Studie 2 verwendet wurden. Nichtsdestotrotz könnten die Belastbarkeit und die Bedeutung der Befunde von Studie 1 und 2 durch Replikationen anhand solcher Stichproben in großem Maße verstärkt werden. Damit wäre auch der Limitation Rechnung getragen, dass es sich bei den in den Studien 1-3 untersuchten Stichproben um Gelegenheits- und nicht um Zufallsstichproben handelt, auch wenn es sich bei den untersuchten Schulen um „typische“ öffentliche Schulen handelt und nicht anzunehmen ist, dass sie oder ihre Schülerschaft sich systematisch von anderen Schulen der jeweiligen Schulform unterscheiden.

Weitere Einschränkungen betreffen das Design der Studien. Studie 1 und Studie 2 konnten wertvolle Erkenntnisse zum Verständnis der Zusammenhänge zwischen verschiedenen Variablen und dem differenziellen Schulerfolg von Mädchen und Jungen liefern. Sie basieren jedoch auf korrelativen Querschnittsdaten, die keine Aussagen über

kausale Wirkungen zwischen den untersuchten Variablen erlauben. Aus diesem Grund wäre es sinnvoll, die in Studie 1 und 2 aufgezeigten Zusammenhänge mithilfe von Längsschnittstudien zu untermauern und vermutete Kausalitäten mit Experimenten gezielt zu überprüfen. Studie 3 hingegen ist eine experimentelle Vignettenstudie, in der die Lehrkräfte zufällig den vier Vignetten zugeteilt wurden. Die damit verbundene relativ hohe interne Validität geht jedoch mit einer geringeren externen Validität einher. Studie 3 weist im Rahmen einer Vignettenstudie einen potentiellen Wirkmechanismus nach, aber eine Ergänzung durch weitere Feldstudien wäre nötig, um zu prüfen, wie extern valide und beständig die gefundenen Ergebnisse tatsächlich sind, da unzutreffende erste Eindrücke und Erwartungen durch den täglichen Kontakt der Lehrkräfte mit den Schülerinnen und Schülern akkurater werden sollten (siehe Raudenbush, 1984).

#### **6.4 Implikationen**

Die Ergebnisse der drei Studien weisen darauf hin, dass Geschlechterstereotype und die geschlechtsbezogene Identität und Identitätsentwicklung Jugendlicher zur Erklärung von Geschlechtsunterschieden im Schulerfolg beitragen können. Des Weiteren verdeutlichen die Ergebnisse, dass es wichtig ist, zwischen den Fächern Deutsch und Mathematik zu unterscheiden und dass Befunde innerhalb eines Faches nicht ohne weiteres auf ein anderes übertragbar sind. Wie im vorigen Abschnitt diskutiert, könnte die zukünftige Forschung zu diesem Thema von Längsschnittstudien oder Experimenten mit repräsentativen Daten und verschiedenen Indikatoren von Schulerfolg profitieren. In den folgenden Abschnitten wird vorgestellt, welche weiteren inhaltlichen Fragestellungen für die zukünftige Forschung und welche Vorschläge für die Praxis im Klassenzimmer aus den Ergebnissen abgeleitet werden können.

### 6.4.1 Forschung

Obwohl sich in expliziten Maßen viele indirekte Hinweise auf eine deutlich stereotypisierte Wahrnehmung von Schule als feminin zeigten (siehe Kapitel 3.2.3), wurde in Studie 2 mithilfe des GNAT nur eine leichte Stereotypisierung von Schule als feminin bei den Jugendlichen nachgewiesen. Dies steht einerseits im beruhigenden Kontrast zur reißerischen medialen Darstellung (zusammenfassend z.B. Latsch & Hannover, 2014). Andererseits wirft es für die weitere Forschung die Frage auf, worauf sich dieser vermeintliche Widerspruch zurückführen lässt. Eine mögliche, noch zu prüfende Erklärung stellt die Erfassung des Stereotyps in Studie 2 mit einem impliziten, also nonreaktiven Messverfahren dar. Eine implizite Erfassung von Stereotypen hat sich in vielen Fällen bewährt, ganz besonders, wenn Aspekte untersucht werden sollten, die als sozial unerwünscht gelten und deren Zustimmung den relevanten Normen widerspricht (z.B. Greenwald, Poehlman, Uhlmann & Banaji, 2009). Bestünde aber in manchen Peergroups die Norm, Schule als feminin „abzuwerten“ und sich als Junge nicht (öffentlich sichtbar) anzustrengen, sollte sich mithilfe expliziter Messverfahren, z.B. Fragebögen, eine stärkere Stereotypisierung von Schule und schulischem Lernen aufzeigen lassen als in impliziten Messverfahren wie dem GNAT. Qualitative Interviewstudien deuten darauf hin, dass dies zumindest für die Gruppe der „lads“ zutrifft (z.B. Jackson, 2002, 2003; Jackson & Dempster, 2009). Um ein tieferes Verständnis der Geschlechtsstereotypisierung von Schule und des vermeintlichen Widerspruchs zur öffentlichen Wahrnehmung zu gewinnen, wäre es folglich sinnvoll, in zukünftigen Studien die Geschlechtsstereotypisierung von Schule und Lernen sowohl implizit als auch explizit zu erfassen. Denn wenn manche Jungen die öffentliche Distanzierung von Schule und schulischem Engagement als Strategie zur Selbstdarstellung und Fremdtäuschung nutzen, weil so eine Distanzierung in ihrer Peergroup angesehen ist (z.B. Jackson & Dempster, 2009), sollten diese Jungen Schule auf expliziten Maßen als femininer einschätzen als in impliziten Messverfahren.

Um nähere Einblicke in die Prozesse zu gewinnen, die dem in Studie 2 aufgezeigten Zusammenhang mit Schulerfolg zu Grunde liegen, wäre es außerdem nützlich, motivationale Variablen und Angaben zum Lernverhalten hinsichtlich ihrer ihnen vom IIRM und der IBMT angedachten vermittelnden Rolle zu untersuchen. Dabei kann *auch* auf das Eccles et al.-Modell als theoretischen Rahmen zurückgegriffen werden, denn die Bedeutung der eigenen Identität und die Identitätskongruenz schulischer Angebote für Motivation und Schulerfolg ist nicht nur im IIRM und in der IBMT verankert, sondern als „Allgemeine Ziele und Selbstbild des Kindes“ auch im Eccles et al.-Modell (z.B. Eccles et al., 1983; Wigfield & Eccles, 2000). So sollte den schulischen Angeboten, mit denen es möglich ist, zentrale Aspekte der eigenen Identität auszudrücken oder zu bestätigen, theoretisch eine höhere *persönliche Wichtigkeit* zugeschrieben werden, als denjenigen Angeboten, die diese Möglichkeit nicht bieten, da sie als weniger passend zur eigenen Identität erlebt werden (Eccles, 2007; Wigfield & Eccles, 1992). Ebenso sollten dem Eccles et al.-Modell zufolge die schulischen Angebote als besonders *nützlich* erlebt werden, die dabei helfen, Fähigkeiten zu entwickeln, die für das zukünftige Selbst zentral sind (z.B. Wigfield & Eccles, 1992). Umgekehrt lässt sich vermuten, dass die Bearbeitung und Nutzung identitätsinkongruenter schulischer Angebote mit höheren *Kosten* einhergeht als die identitätskongruenter. Für die weitere Forschung leitet sich daraus die Frage ab, ob die in Studie 2 nachgewiesene Stereotypisierung von Schule und Lernen als feminin erklären kann, warum Jungen Schule insgesamt einen geringeren Wert zuschreiben (siehe Kessels & Steinmayr, 2013a; Steinmayr & Spinath, 2010), und ob die möglicherweise hohen Kosten, die aufgrund der Geschlechtsstereotypisierung für manche Jungen mit schulischem Engagement verbunden sein könnten, erklären können, warum Jungen ein ungünstigeres Lernverhalten und weniger Engagement berichten als Mädchen (siehe Lam et al., 2012; Trautwein, Lüdtke, Kastens & Köller, 2006). In diesem Fall sollten schul- und lernbezogene Werte und Kosten und das Lernverhalten den Zusammenhang zwischen der Geschlechtsstereotypisierung von Schule und dem Schulerfolg von Jungen mediiieren. Um

diese Fragen empirisch untersuchen zu können, erscheint es jedoch nötig, die bisherigen im deutschen Sprachraum verwendeten Skalen zur Erfassung schulischer Werte und Kosten weiterzuentwickeln, da sie die theoretisch enthaltenden Identitätsaspekte nicht ausreichend berücksichtigen (z.B. Gaspard et al., 2014; Steinmayr & Spinath, 2010; für eine englischsprachige Ausnahme siehe Conley, 2012).

Des Weiteren stimmen das Eccles et al.-Modell, das IIRM und die IBMT darin überein, dass der Identität bzw. verschiedenen Identitätsaspekten von Schülerinnen und Schülern nicht in jeder Situation die gleiche Bedeutung für ihre Motivation, Lernverhalten und Schulerfolg zukommt, was in den Studien 1-3 nicht berücksichtigt worden ist. So ist gemäß dem IIRM (z.B. Kessels & Hannover, 2004, 2006) und der IBMT (z.B. Oyserman & Destin, 2010) die Passung von schulischen Anforderungen und Aufgaben zu demjenigen Identitätsaspekt besonders handlungsleitend, der aufgrund situativer Merkmale gerade salient ist (siehe auch Hannover, 1997; Linville & Carlston, 1994; Markus & Wurf, 1987). Gemäß dem Eccles et al.-Modell variieren auch die persönliche Wichtigkeit und Nützlichkeit einer Aufgabe in Abhängigkeit davon, welcher Identitätsaspekt gerade besonders salient ist (z.B. Eccles, 2007). Auch wenn in der Adoleszenz von einer permanent hohen Salienz von Geschlecht ausgegangen werden kann (z.B. Finkenauer et al., 2002; Kessels & Hannover, 2004), sind stärkere Zusammenhänge zwischen der Stereotypisierung von Schule und der Motivation und dem Lernverhalten von Jugendlichen zu erwarten, wenn die Salienz des Faktors Geschlecht erhöht und damit die geschlechtsbezogene Identität besonders zugänglich ist. Gelänge es hingegen die Salienz von Geschlecht in der Schule zu reduzieren, sollte einer Stereotypisierung von Schule als feminin eine geringere Bedeutung zukommen. Eine vergleichbare Moderatorfunktion der Salienz von Geschlecht ist bereits für die Wirkung anderer Stereotype nachgewiesen worden (z.B. Bigler, 1995; Hilliard & Liben, 2010; Kessels & Hannover, 2008). Eine experimentelle Prüfung dieser Moderatorfunktion für die Bedeutung der Stereotypisierung von Schule in zukünftigen Studien erscheint nicht nur aus theoretischen,

sondern auch aus praktischen Gründen sinnvoll, da sich aus den Ergebnissen ggf. auch konkrete Handlungsempfehlungen für Lehrkräfte ableiten ließen (siehe Kapitel 6.4.2).

Außerdem stellt sich die Frage, was die interindividuellen Unterschiede in dem Ausmaß der wahrgenommenen Stereotypisierung von Schule und schulischem Lernen erklärt. Hier erscheint es besonders reizvoll, die Rolle der jeweiligen Peergroup und ihrer Normen näher zu betrachten. Denn Studien zu Klassen- oder Schulkompositionseffekten deuten darauf hin, dass die Zusammensetzung der Peergroup besonders bei Jungen Einfluss darauf hat, welche schulbezogenen Einstellungen und welches schulisches Verhalten als angemessen gelten und gezeigt werden (z.B. Demanet, Vanderwegen, Vermeersch & van Houtte, 2013; Legewie & DiPrete, 2012) und wie groß etwaige Geschlechtsunterschiede im Schulerfolg zu Ungunsten der Jungen ausfallen (z.B. Legewie & DiPrete, 2012; van de gaer, Pustjens, Damme & Munter, 2006; van de gaer, Pustjens, van Damme & Munter, 2007; van Houtte, 2004).

In den Studien der vorliegenden Arbeit wurde die Bedeutung von Geschlechterstereotypen, der geschlechtsbezogenen Identität und des untersuchten Fachs für die Erklärung des differenziellen Schulerfolgs von Mädchen und Jungen getrennt an zwei Schüler- und an einer Lehrkräftestichprobe untersucht. Im schulischen Alltag interagieren beide Gruppen jedoch miteinander und ihre Kognitionen und ihr Verhalten stehen in einem reziproken Zusammenhang, wie nicht nur Studien zur Wirkung von leistungsbezogenen Lehrererwartungen (z.B. Brattesani, Weinstein & Marshall, 1984; Harris & Rosenthal, 1985; Jussim & Harber, 2005), sondern auch Studien zu verhaltensbezogenen Lehrererwartungen gezeigt haben (Demanet & van Houtte, 2012; Skinner & Belmont, 1993). Die reziproken Zusammenhänge können eine Negativspirale bilden, wie an zwei Beispielen auf Grundlage bisheriger Studien veranschaulicht wird. Die erste Negativspirale könnte durch das reziproke Verhältnis von Schüler- und Lehrermotivation und -verhalten entstehen. Denn Demanet und van Houtte (2012) zeigten, dass wenn Lehrkräfte einen negativen Eindruck bezüglich des Lern- und Sozialverhaltens und der Motivation ihrer Schülerinnen und Schülern hatten, die

Schülerinnen und Schüler sich weniger unterstützt fühlten und mit vermehrtem Störverhalten und weniger Engagement und Mitarbeit reagierten. Weitere Studien zeigten, dass Lehrkräfte auf Störverhalten und geringes Engagement wiederum mit weniger Unterstützung und weniger motivationsfördernden Angeboten reagierten, welches wiederum mit einem ungünstigerem Lernverhalten und negativeren Emotionen der Schülerinnen und Schüler einherging (z.B. Rimm-Kaufman, Baroody, Larsen, Curby & Abry, 2015; Skinner & Belmont, 1993). Daraus lässt sich die Vermutung ableiten, dass geringere Erwartungen und negativere Einstellungen von Lehrkräften gegenüber manchen Schülerinnen und Schülern, beispielsweise – wie in Studie 3 gezeigt – aufgrund von Stereotypenaktivierung, sich ungünstig auf das Verhalten der Lehrkräfte gegenüber diesen Jugendlichen auswirken und die betroffenen Jugendlichen sich als Konsequenz weniger für die Schule anstrengen und weniger zugehörig fühlen. Bei manchen Jungen könnte diese Negativspirale zu dem Verständnis und Gefühl führen, Schule und schulisches Lernen seien eher etwas für Mädchen als für Jungen (siehe Studie 2), welches wiederum mit einer weiteren Reduktion schulischer Anstrengungsbereitschaft und – wie in Studie 2 gezeigt – schlechteren Noten einhergehen könnte. Denn ein höheres Zugehörigkeitsgefühl zur Schule sagt prospektiv bessere Noten und eine höhere Wertschätzung von Schule vorher (z.B. Gillen-O'Neel & Fuligni, 2013; Good, Rattan & Dweck, 2012).

Alternativ wäre es auch vorstellbar, dass die maskulinen Jungen, die Schule als feminin stereotypisiert wahrnehmen, ein größeres Bedürfnis haben, ihre Maskulinität in der Schule zu demonstrieren – möglicherweise besonders im feminin stereotypisierten Fach Deutsch. Die in Studie 3 aufgezeigten Effekte von Gender Enactment auf das den Jugendlichen zugeschriebene Lernverhalten sowie die Hinweise, dass Lehrkräfte des muttersprachlichen Fachs schulische Anstrengung stärker mitbewerten als Mathematiklehrkräfte (z.B. McMillan, 2001; Resh, 2009), könnten sich dann wiederum in schlechteren Deutsch- aber nicht schlechteren Mathematiknoten der Jungen manifestieren, die

Schule als feminin wahrnehmen und sich selbst sehr maskulin einschätzen (siehe Studie 2). Diese negativen Leistungsrückmeldungen könnten bei den betroffenen Jungen die Wahrnehmung von Schule als feminin erneut verstärken.

Für die zukünftige Forschung stellt somit die Analyse der Interaktionsprozesse zwischen Schülerinnen, Schülern und Lehrkräften in Längsschnittstudien eine weitere, herausfordernde Aufgabe dar. Besonders aufschlussreich wäre hier auch die Ergänzung der Selbstberichte zum Lernverhalten und Engagement von Schülerinnen, Schülern und Lehrkräften um objektivere Beobachtungsmaße, um die Bedeutung des Verhaltens an sich von der Bedeutung der Wahrnehmung und Interpretation des Verhaltens trennen zu können.

#### **6.4.2 Praxis**

Die Ergebnisse der Studien 2 und 3 zeigten auf, dass die stereotype Verknüpfung von Schule und schulischem Lernverhalten mit Mädchen und feminin vor dem Hintergrund der Identitätsentwicklung eine weitere wichtige Erklärung für den geringeren Schulerfolg mancher Jungen darstellt. Daraus lassen sich verschiedene Vorschläge für das Klassenzimmer ableiten. Diese wenden sich an Lehrkräfte und betreffen (1) die Bedeutung von Stereotypen und (2) die Bedeutung der geschlechtsbezogenen Identität und Identitätsentwicklung Jugendlicher.

Die Zugänglichkeit und damit der potentielle Einfluss eines Stereotyps sind theoretisch dann gering, wenn die dem Stereotyp zu Grunde liegende Kategorie wenig salient ist (z.B. Kessels & Hannover, 2008; Shih et al., 1999). Aus diesem Befund lässt sich zwar nicht konkret ableiten, wie Lehrkräfte sich verhalten sollen, damit soziale Kategorien möglichst wenig salient sind, aber zumindest, wie sie sich *nicht* verhalten sollten. Beispielsweise empfiehlt es sich, das Geschlecht der Schülerinnen und Schüler nicht zur Klassenorganisation zu verwenden und die Betitelung von Aufgaben, Tätigkeiten, Verhaltensweisen u.a. als Jungen- und Mädchenaufgaben etc. strikt zu vermeiden, da dies Geschlechterstereotype bei



den Schülerinnen und Schülern aktiviert (z.B. Hilliard & Liben, 2010). Außerdem könnten Lehrkräfte darauf achten, dass Tätigkeiten, wie beispielsweise die Klassenbuchführung, das Schmücken des Klassenzimmers oder das Umstellen von schweren Tischen für Gruppenarbeiten, von Jungen und Mädchen in gleichem Ausmaß übernommen werden, um eine mögliche Geschlechtsstereotypisierung dieser Tätigkeiten abzuschwächen und entsprechenden Geschlechtsrollenerwartungen entgegenzuwirken. Gleichzeitig erscheint es wichtig, sowohl für Lehrkräfte als auch für Schülerinnen und Schüler Anregungen zu schaffen, die eigenen Geschlechterstereotype in dem Maße, in dem es möglich ist, kritisch zu reflektieren, um ihren potentiellen Einfluss zu reduzieren. Dazu könnte z.B. auf meta-analytische Befunde zur Gender Similarities Hypothese (Hyde, 2005) zurückgegriffen werden, welche u.a. belegen, dass die Unterschiede zwischen den Geschlechtergruppen mit wenigen Ausnahmen klein sind, vor allem kleiner sind als die interindividuellen Unterschiede innerhalb einer Geschlechtergruppe und z.B. in Abhängigkeit vom sozialen Kontext in ihrer Größe variieren. Außerdem könnten konkrete Gegenbeispiele helfen, die grobe Verallgemeinerung, die Stereotypen anhaftet, und ihre Gültigkeit für das Individuum zu hinterfragen und die Ausprägung des Stereotyps abzuschwächen (z.B. Dasgupta & Asgari, 2004; Elmore & Oyserman, 2012; Kessels et al., 2006; Lenton, Bruder & Sedikides, 2009). Vor dem Hintergrund der Studien der vorliegenden Arbeit wäre das Ziel solcher Reflexionsprozesse, dass Lehrkräfte, Schülerinnen und Schüler Schule, schulische Anstrengung und Schulerfolg allgemein als zu Jungen und Mädchen gleich gut passend wahrnehmen und Schülerinnen *und* Schüler sie als kongruent zu ihrer geschlechtsbezogenen Identität erleben können.

Lehrkräfte hingegen könnte das Wissen um die Aufgabe der Jugendliche, eine geschlechtsbezogene Identität zu entwickeln, entlasten. Das Bewusstsein, dass schulisches Verhalten von Jugendlichen auch zur Identitätsentwicklung genutzt wird, könnte beispielsweise zu einem entspannteren Umgang mit ablehnendem Schülerverhalten führen.

Wenn eine zur Schau gestellte Ablehnung von Schule oder dem Fach Physik nicht auf mangelnde eigene didaktische Fähigkeiten attribuiert, sondern einfach als Mittel zur Identitätsentwicklung Jugendlicher verstanden würde, sollte dies nach Weiner (1985) den Selbstwert der Lehrkräfte schützen und ihre Motivation und Selbstwirksamkeitserwartung erhöhen (bzw. nicht reduzieren). Ähnliches war bereits global für die Attribution von Stress bei Lehrkräften nachgewiesen worden (Wang, Hall & Rahimi, 2015). Von motivierten Lehrkräften mit hohen Selbstwirksamkeitserwartungen profitieren wiederum auch die Schülerinnen und Schüler (z.B. Caprara, Barbaranelli, Steca & Malone, 2006; Demanet & van Houtte, 2012).

## **6.5 Fazit**

Ziel der vorliegenden Arbeit war es, zur Erklärung des differenziellen Schulerfolgs von Mädchen und Jungen in Deutschland beizutragen. Auf Grundlage des Eccles et al.-Modells, des IIRM und der IBMT sowie des Kontinuum-Modells der Eindrucksbildung wurden drei separate Erklärungen entwickelt und in drei empirischen Studien überprüft. Dazu wurden Strukturgleichungsmodelle, ein computerbasiertes, implizites Assoziationsmaß, der GNAT, sowie eine experimentelle Vignettenstudie genutzt. Die Ergebnisse machen vier bedeutende Punkte sichtbar. (1) Verhaltensbezogene Geschlechterstereotype und eine geschlechtsstereotypisierte Wahrnehmung von Schule stellen eine weitere Erklärung für den differenziellen Schulerfolg von Jungen und Mädchen dar. Entgegen der medialen Darstellung wird Schule aber keinesfalls als ein durch und durch feminisierter Raum wahrgenommen. Überfachliche und fachspezifische Geschlechterstereotype interagieren und sollten beide berücksichtigt werden. (2) Die geschlechtsbezogene Identität von Jugendlichen und ihre Entwicklung während der Adoleszenz kann sowohl mit dem Schulerfolg von Jungen als auch mit dem von Mädchen im Konflikt stehen, je nachdem wie die eigene Geschlechtszugehörigkeit und das eigene geschlechtsbezogene Selbstkonzept ausgestaltet

sind und werden. Allerdings sind mögliche Konflikte für Jungen gegenwärtig wahrscheinlicher als für Mädchen. (3) Ergebnisse zur Erklärung von Geschlechtsunterschieden im Schulerfolg lassen sich nicht ohne weiteres von einem Fach auf ein anderes übertragen. Eine wahrscheinliche Erklärung stellt die unterschiedliche Geschlechtsstereotypisierung von Schulfächern dar. (4) Auch hinter auf den ersten Blick besseren Noten von Jungen kann sich angesichts der Geschlechtsunterschiede in der Bildungsbeteiligung ein – relativ zu ihrem Potential gesehen – geringerer Schulerfolg der Jungen verbergen. Aus diesem Grund sind Schulnoten als alleiniger Indikator für Schulerfolg weniger aussagekräftig, als wenn zusätzlich das kognitive, motivationale oder soziale Potential der Schülerinnen und Schüler berücksichtigt wird.

Aus den Ergebnissen lassen sich vielfältige Anregungen für die weitere Forschung und die schulische Praxis ableiten. Gemeinsam stellt sich für beide Bereiche die Aufgabe, ein Bewusstsein für die Bedeutung der Identität und Geschlechterstereotypen für das Verhalten und die Kognitionen von Jungen, Mädchen und Lehrkräften zu schaffen, um dem Ideal der Bildungsgerechtigkeit ein Stück näher zu kommen.



## Literaturverzeichnis II

- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development, 66*, 1072–1087.
- Brattesani, K. A., Weinstein, R. S. & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology, 76*, 236–247.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*, 123–142.
- Caprara, G. V., Barbaranelli, C., Steca, P. & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44*, 473–490.
- Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*, 32–47.
- Dasgupta, N. & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642–658.
- Demanet, J. & van Houtte, M. (2012). Teachers' attitudes and students' opposition. School misconduct as a reaction to teachers' diminished effort and affect. *Teaching and Teacher Education, 28*, 860–869.
- Demanet, J., Vanderwegen, P., Vermeersch, H. & van Houtte, M. (2013). Unravelling gender composition effects on rule-breaking at school: a focus on study attitudes. *Gender and Education, 25*, 466–485.
- DiPrete, T. A. & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research, 41*, 1–15.
- Downey, D. B. & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly, 46*, 299–321.
- Duckworth, A. L., Quinn, P. D. & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology, 104*, 439–451.
- Duckworth, A. L. & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198–208.
- Dumont, H., Trautwein, U., Lüdtke, O., Neumann, M., Niggli, A. & Schnyder, I. (2012). Does parental homework involvement mediate the relationship between family background and educational outcomes? *Contemporary Educational Psychology, 37*, 55–69.

- Eccles, J. S. (2007). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Hrsg.), *Handbook of competence and motivation* (S. 105–121). New York, NY: Guilford Press.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives. Psychological and sociological approaches* (S. 76–146). San Francisco, CA: W.H. Freeman.
- Eckes, T. & Trautner, H. M. (2000). Developmental social psychology of gender: An integrative framework. In T. Eckes & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 3–32). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Elmore, K. C. & Oyserman, D. (2012). If 'we' can succeed, 'I' can too: Identity-based motivation and gender in the classroom. *Contemporary Educational Psychology*, 37, 176–185.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. New York, NY: Norton.
- Fiedler, K. & Bless, H. (2003). Soziale Kognition. In W. Stroebe, K. Jonas & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (4. Aufl., S. 125–164). Heidelberg: Springer.
- Finkenauer, C., Engels, R. C., Meeus, W. & Oosterwegel, A. (2002). Self and identity in early adolescence: The pains and gains of knowing who and what you are. In T. M. Brinthaup & R. P. Lipka (Hrsg.), *Understanding early adolescent self and identity. Applications and interventions* (S. 25–56). Albany, NY: State University of New York Press.
- Fiske, S. T. & Neuberg, S. L. (1990). A continuum of impression formation, from category based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74.
- Frome, P. M. & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74, 435–452.
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U. et al. (2014). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, Advance online publication.
- Gillen-O'Neel, C. & Fuligni, A. (2013). A longitudinal study of school belonging and academic motivation across high school. *Child Development*, 84, 678–692.
- Good, C., Rattan, A. & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102, 700–717.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L. & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97, 17–41.
- Hannover, B. (1997). *Das dynamische Selbst. Die Kontextabhängigkeit selbstbezogenen Wissens*. Bern: Verlag Hans Huber.
- 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.
- Harris, M. J. & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, 97, 363–386.

- Hartley, B. L. & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development*, 84, 1716–1733.
- Havighurst, R. J. (1956). Research on the developmental-task concept. *The School Review*, 64, 215–223.
- Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 93–111.
- Herbert, J. & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, 26, 276–295.
- Hilliard, L. J. & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, 81, 1787–1798.
- Hosenfeld, I., Köller, O. & Baumert, J. (1999). Why sex differences in mathematics achievement disappear in German secondary schools: A reanalysis of the German TIMSS-data. *Studies in Educational Evaluation*, 25, 143–161.
- Hülür, G., Wilhelm, O. & Robitzsch, A. (2011). Multivariate Veränderungsmodelle für Schulnoten und Schülerleistungen in Deutsch und Mathematik. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 173–185.
- Hyde, J. S. (2005). The Gender Similarities Hypothesis. *American Psychologist*, 60, 581–592.
- Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy. *Gender and Education*, 14, 37–50.
- Jackson, C. (2003). Motives for 'laddishness' at school: Fear of failure and fear of the 'feminine'. *British Educational Research Journal*, 29, 583–598.
- Jackson, C. & Dempster, S. (2009). 'I sat back on my computer ... with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, 18, 341–356.
- Jussim, L. & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131–155.
- Kaiser, J., Retelsdorf, J., Südkamp, A. & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, 28, 73–84.
- Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M. & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, 42, 11–26.
- 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 Qualitätsentwicklung* (S. 398–412). Münster: Waxmann.
- Kessels, U. & Hannover, B. (2006). Zum Einfluss des Image von mathematisch-naturwissenschaftlichen Schulfächern auf die schulische Interessensentwicklung. In M. Prenzel (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 350–369). Münster: Waxmann.

- Kessels, U. & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. *British Journal of Educational Psychology*, 78, 273–289.
- Kessels, U., Heyder, A., Latsch, M. & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, 56, 219–228.
- Kessels, U., Rau, M. & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 76, 761–780.
- Kessels, U. & Steinmayr, R. (2013a). Der subjektive Wert von Schule in Abhängigkeit vom verbalen und mathematischen Selbstkonzept. *Zeitschrift für Pädagogische Psychologie*, 27, 105–113.
- Kessels, U. & Steinmayr, R. (2013b). 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.
- Kessels, U., Warner, L. M., Holle, J. & Hannover, B. (2008). Identitätsbedrohung durch positives schulisches Leistungsfeedback. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40, 22–31.
- Lam, S.-f., Jimerson, S., Kikas, E., Cefai, C., Veiga, F. H., Nelson, B. et al. (2012). Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of School Psychology*, 50, 77–94.
- 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.
- Legewie, J. & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review*, 77, 463–485.
- Lehmann, R. H., Peek, R. & Gänsfuß, R. (1997). *Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen*. Hamburg: Behörde für Schule, Jugend und Berufsbildung.
- Lenton, A. P., Bruder, M. & Sedikides, C. (2009). A meta-analysis on the malleability of automatic gender stereotypes. *Psychology of Women Quarterly*, 33, 183–196.
- Leslie, S.-J., Cimpian, A., Meyer, M. & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347, 262–265.
- Linville, P. W. & Carlston, D. E. (1994). Social cognition of the self. In P. G. Devine, D. L. Hamilton & T. M. Ostrom (Hrsg.), *Social cognition: Impact on social psychology* (S. 143–193). San Diego, CA: Academic Press.
- Lyng, S. T. (2009). Is there more to "antischoolishness" than masculinity?: On multiple student styles, gender, and educational self-exclusion in secondary school. *Men and Masculinities*, 11, 462–487.
- Maaz, K. & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems. Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In Bundesministerium für Bildung und Forschung (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 151–180). Berlin: Bundesministerium für Bildung und Forschung.



- MacKinnon, D. P., Krull, J. L. & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*, 173-181.
- Markus, H. & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38*, 299-337.
- McDermott, P. A., Mordell, M. & Stoltzfus, J. C. (2001). The organization of student performance in American schools: Discipline, motivation, verbal learning, nonverbal learning. *Journal of Educational Psychology, 93*, 65-76.
- McKenney, S. J. & Bigler, R. S. (2014). High heels, low grades: Internalized sexualization and academic orientation among adolescent girls. *Journal of Research on Adolescence*, Advance online publication.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*, 20-32.
- McMillan, J. H., Myran, S. & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research, 95*, 203-213.
- Nguyen, H.-H. D. & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*, 1314-1334.
- Nosek, B. A. & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition, 19*, 625-666.
- Nosek, B. A., Banaji, M. R. & Greenwald, A. G. (2002). Math = male, me = female, therefore math  $\neq$  me. *Journal of Personality and Social Psychology, 83*, 44-59.
- Nosek, B. A. & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal, 48*, 1125-1156.
- Oyserman, D. & Destin, M. (2010). Identity-Based Motivation: Implications for intervention. *The Counseling Psychologist, 38*, 1001-1043.
- Plante, I., de la Sablonnière, R., Aronson, J. M. & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology, 38*, 225-235.
- Randall, J. & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education, 26*, 1372-1380.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology, 76*, 85-97.
- Resh, N. (2009). Justice in grades allocation: teachers' perspective. *Social Psychology of Education, 12*, 315-325.
- Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A. A., Curby, T. W. & Abry, T. (2015). To what extent do teacher-student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning? *Journal of Educational Psychology, 107*, 170-185.
- Shih, M., Pittinsky, T. L. & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*, 80-83.
- Skinner, E. A. & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*, 571-581.

- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4–28.
- Stanat, P. & Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 249–269). Opladen: Leske + Budrich.
- Statistisches Bundesamt. (2013). *Bildung und Kultur. Allgemeinbildende Schulen Schuljahr 2012/2013*. Wiesbaden: Statistisches Bundesamt.
- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811.
- Steffens, M. C., 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.
- Steinmayr, R. & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica*, 56, 195–211.
- Trautwein, U., Lüdtke, O., Kastens, C. & Köller, O. (2006). Effort on homework in grades 5–9: Development, motivational antecedents, and the association with effort on classwork. *Child Development*, 77, 1094–1111.
- van de gaer, E., Pustjens, H., Damme, J. & Munter, A. (2006). The gender gap in language achievement: The role of school-related attitudes of class groups. *Sex Roles*, 55, 397–408.
- van de gaer, E., Pustjens, H., van Damme, J. & Munter, A. de. (2007). Impact of attitudes of peers on language achievement: Gender differences. *The Journal of Educational Research*, 101, 78–90.
- 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.
- Verordnung über die Schularten und Bildungsgänge der Sekundarstufe I. (2010). Sekundarstufe I-Verordnung. Sek I-VO. Verfügbar unter <http://gesetze.berlin.de/jportal/portal/t/dut/page/bsbeprod.psml?action=controls.jw.PrintOrSaveDocumentContent&case=save>
- Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204.
- Wang, H., Hall, N. C. & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, 47, 120–130.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.
- Wigfield, A. & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12, 265–310.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.

Ziegenspeck, J. W. (1999). *Handbuch Zensur und Zeugnis in der Schule*. Bad Heilbrunn: Klinkhardt.



## **Erklärung**

Hiermit versichere ich, dass ich die Dissertation „Mehr als Erwartungen und Werte: Zur Rolle von Stereotypen für den differenziellen Schulerfolg von Mädchen und Jungen“ selbstständig verfasst habe. Sämtliche Hilfsmittel, die ich verwendet habe, sind angegeben. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, im April 2015

---



## **Lebenslauf**

Der Lebenslauf ist aus Gründen des Datenschutzes nicht enthalten.

