

# Kohortentrends in schulfachbezogenen Selbstkonzepten und Interessen bei Mädchen und Jungen

Rebecca Schneider<sup>1</sup> , Sarah Gentrup<sup>2</sup> , Malte Jansen<sup>1,3</sup>  und Petra Stanat<sup>1</sup> 

<sup>1</sup>Institut zur Qualitätsentwicklung im Bildungswesen e.V. an der Humboldt-Universität zu Berlin, Deutschland

<sup>2</sup>Humboldt-Universität zu Berlin, Deutschland

<sup>3</sup>Zentrum für internationale Vergleichsstudien (ZIB), München, Deutschland

**Zusammenfassung:** Schulfachbezogene Selbstkonzepte und Interessen gehören zu den wichtigsten motivationalen Konstrukten in der Pädagogischen Psychologie und zeigen typischerweise stereotype Geschlechtsunterschiede: Während Jungen in Mathematik und den Naturwissenschaften im Mittel ein höheres Selbstkonzept und Interesse aufweisen, berichten Mädchen höhere Werte in sprachlichen Fächern. Erste empirische Ergebnisse von wiederholt durchgeführten Studien des Bildungsmonitorings weisen auf einen leichten Rückgang im Selbstkonzept und Interesse in Deutschland über den Verlauf der untersuchten Kohorten hin. Jüngste Befunde zeigen diesen Rückgang insbesondere für Jungen im MINT-Bereich. Trotz substantieller Zusammenhänge zwischen motivationalen Merkmalen und schulischen Leistungen wurde bisher nicht untersucht, inwiefern diese Motivationsrückgänge auf Veränderungen in schulischen Leistungen zurückgeführt werden können. Der vorliegende Beitrag untersucht deshalb, (a) ob die Kohortentrends in den fachbezogenen Selbstkonzepten und Interessen für Jungen und Mädchen je nach Fach unterschiedlich ausfallen und insbesondere, (b) ob sich etwaige (geschlechtsspezifische) Trends in der fachbezogenen Motivation auf Unterschiede in Testleistungen und Noten zwischen den Erhebungszeitpunkten zurückführen lassen. Auf Basis der IQB-Bildungstrendstudien wurden Daten von je etwa 25 000 Neuntklässlerinnen und Neuntklässlern in den Jahren 2012 und 2018 (Fächer Mathematik, Physik, Chemie und Biologie) bzw. von je etwa 35 000 Neuntklässlerinnen und Neuntklässlern in den Jahren 2009 und 2015 (Fach Deutsch) ausgewertet. Während sich für Mädchen in Mathematik und den naturwissenschaftlichen Fächern keine oder geringfügig positive Veränderungen zwischen den Kohorten zeigten ( $-0.05 \leq d \leq 0.07$ ), fanden sich für Jungen zumeist (leichte) Rückgänge in den Selbstkonzepten und Interessen ( $-0.31 \leq d \leq -0.08$ ). Insbesondere die Veränderung im Mathematikinteresse bei Jungen war bedeutsam ( $d = -0.31$ ). Im Fach Deutsch zeigte sich ein gegensätzliches Bild: Bei beiden Geschlechtern stiegen die Selbstkonzeptwerte im Kohortentrend geringfügig an (Mädchen/Jungen:  $d = 0.07 / 0.06$ ). Diese geschlechtsspezifischen Trends in der Motivation ließen sich nicht bedeutsam auf Testleistungs- und Notenunterschiede zwischen den Kohorten zurückführen. Die Ergebnisse werden im Hinblick auf verschiedene Erklärungsansätze für (geschlechtsspezifische) Kohortentrends in schulischer Motivation diskutiert.

**Schlüsselwörter:** Selbstkonzept, Interesse, schulische Leistung, Sekundarstufe, Kohortenvergleich

## Cohort trends for school subject-specific self-concepts and interests for girls and boys

**Abstract:** Academic self-concepts and interests are among the most important motivational constructs in educational psychology. Gender differences in these constructs are in accordance with prevailing stereotypes: Whereas boys typically report higher self-concepts and interests in mathematics and science, girls show higher self-concepts and interests in verbal subjects. So far, the few studies examining temporal changes in these variables between student cohorts indicate a decline in Germany. A recent study showed that this decline in student motivation might be more pronounced for boys, particularly in mathematics and science. Although motivational constructs and academic achievement are related, differences in academic achievement between cohorts have not been explored as a possible explanation for cohort trends in self-concepts and interests. Therefore, we examined (a) if cohort trends in students' school subject-specific motivation vary by gender and especially (b) if potential (gender-specific) cohort trends in student motivation can be explained by cohort trends in students' academic achievements (test performance, grades). To do so, we analyzed data of the IQB national assessment studies of student achievement with approx. 25,000 ninth-grade students in 2012 and 2018 (mathematics, physics, chemistry, biology) and with approx. 35,000 ninth-grade students in 2009 and 2015 (German), respectively. Comparing cohorts, girls' self-concepts and interests in mathematics and science did not vary substantially or were marginally higher for the second measurement point ( $-0.05 \leq d \leq 0.07$ ). For boys, however, we found significant declines in self-concepts and interests ( $-0.31 \leq d \leq -0.08$ ). The decline in boys' interest in mathematics was substantial ( $d = -0.31$ ). For the subject German, self-concepts increased slightly for girls ( $d = 0.07$ ) and boys ( $d = 0.06$ ). In all examined subjects, these gender-specific trends in motivation remained widely stable after accounting for academic achievement differences between cohorts. The results are discussed in terms of various explanations for (gender-specific) trends in student motivation.

**Keywords:** self-concept, interest, academic achievement, secondary school, cohort comparison

## Theoretischer Hintergrund

Fachbezogene Selbstkonzepte und Interessen von Schülerinnen und Schülern zählen zu den am häufigsten untersuchten motivationalen Konstrukten. Für beide Konstrukte sind Geschlechtsunterschiede gut belegt: Jungen berichten im Mittel höhere Selbstkonzepte und Interessen in Mathematik und in den naturwissenschaftlichen Fächern, Mädchen hingegen weisen im Mittel höhere Ausprägungen in sprachlichen Fächern auf (z.B. Böhme, Sebald, Weirich & Stanat, 2016; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Jansen, Schneider, Schipolowski & Henschel, 2019; Schilling, Sparfeldt & Rost, 2006; Wilgenbusch & Merrell, 1999). Diese Geschlechtsunterschiede bleiben zumeist auch nach Kontrolle von fachbezogenen Leistungen bestehen (z.B. Jansen, Schroeders & Lütke, 2014; Parker, van Zanden & Parker, 2018).

Wenige Studien haben sich bisher mit Unterschieden in fachbezogenen Selbstkonzepten und Interessen zwischen Kohorten von Schülerinnen und Schülern beschäftigt – also mit der Entwicklung motivationaler Merkmale im Zeitverlauf auf Populationsebene. Erste Ergebnisse von wiederholten Studien des Bildungsmonitorings geben Hinweise auf einen leichten Rückgang im Selbstkonzept und Interesse in Deutschland, d.h. niedrigere Selbstkonzept- und Interessenswerte in späteren Kohorten (Jansen et al. 2019; Jansen, Stäbler, Becker & Neumann, 2017; Schiepe-Tiska & Schmidtner, 2013; Schiepe-Tiska, Simm & Schmidtner, 2016). Jüngste geschlechtsspezifische Analysen im Rahmen des IQB-Bildungstrends 2018 deuten auf Geschlechtsunterschiede in diesem Kohortentrend hin: Während sich die Selbstkonzept- und Interessenswerte der Jungen in Mathematik, Physik und Chemie zwischen 2012 und 2018 verringerten, blieben sie bei Mädchen weitgehend stabil (Jansen et al., 2019).

Anschließend an diese Befunde soll im vorliegenden Beitrag untersucht werden, ob in Mathematik und in den naturwissenschaftlichen Fächern geschlechtsspezifische Kohortenunterschiede in den fachbezogenen Selbstkonzepten und Interessen auf Unterschiede in schulischen Leistungen zwischen Kohorten zurückgeführt werden können. Außerdem soll für die sprachliche Domäne das Selbstkonzept im Fach Deutsch in den Blick genommen werden. Im Fach Deutsch sind uns keine Studien zu geschlechtsspezifischen Kohortentrends bekannt.

## Selbstkonzept und Interesse

Motivationale Merkmale stellen wichtige Zielgrößen des Bildungssystems dar und werden regelmäßig in Schulleistungsstudien zur Beobachtung von Bildungssystemen

untersucht. Zwei zentrale motivationale Konstrukte sind das Selbstkonzept und das Interesse, die sich im Erwartungs-Wert-Modell zur Vorhersage von Bildungsergebnissen (Eccles & Wigfield, 2020) folgendermaßen verorten lassen: Fachbezogene schulische Selbstkonzepte bilden stabile Selbsteinschätzungen der eigenen Fähigkeiten von Schülerinnen und Schülern in einem bestimmten Schulfach ab und sind im Erwartungs-Wert-Modell der Erwartungskomponente zuzuordnen. Das fachbezogene Interesse äußert sich durch anhaltende Wertschätzung und positive Emotionen gegenüber fachlichen Inhalten sowie durch die Beschäftigung mit diesen Inhalten (Schiefele, 2009). Die gleiche Konzeptualisierung wird von den Autorinnen des Erwartungs-Wert-Modells für den intrinsischen Wert genutzt (wird teilweise auch *interest value* genannt; zur Frage der konzeptionellen Gleichsetzung von Interesse und intrinsischem Wert siehe Eccles & Wigfield, 2020), welcher Bestandteil der Wert-Komponente im Erwartungs-Wert-Modell ist.

Die Bedeutung der Erfassung dieser fachbezogenen motivationalen Konstrukte im Rahmen des Bildungsmonitorings liegt insbesondere in ihrem Zusammenhang mit korrespondierenden schulischen Leistungen und bildungsbezogenem Verhalten (z.B. Besuch extracurricularer und außerschulischer Aktivitäten, Kurs- und Berufswahlentscheidungen; z.B. Eccles & Wigfield, 2002; Krapp, 2005; Nagy, Trautwein, Baumert, Köller & Garrett, 2006; Wang & Degol, 2013). Innerhalb eines Schulfachs korrelieren schulische Selbstkonzepte und Interessen typischerweise hoch miteinander (z.B. Eccles & Wigfield, 1995; Rost, Sparfeldt & Schilling, 2007; Trautwein et al., 2012), weisen jedoch differentielle Zusammenhänge mit schulischen Leistungen von Schülerinnen und Schülern auf. So korrelieren fachbezogene Selbstkonzepte mit korrespondierenden schulischen Leistungen meist moderat bis hoch miteinander (Valentine, DuBois & Cooper, 2004), zwischen fachbezogenen Interessen und Leistungen sind die Korrelationen hingegen meist schwach bis moderat ausgeprägt (z.B. Jansen, Lütke & Schroeders, 2016; Nagy et al., 2006; Schiefele, Krapp & Winteler, 1992). Die Korrelationen fallen dabei stärker mit Noten aus als mit Kompetenztestergebnissen (z.B. Jansen et al., 2014, 2016; Marsh, Trautwein, Lütke, Köller & Baumert, 2005), da schulische Motivation und damit zusammenhängende Verhaltensweisen (z.B. Lerngewohnheiten, Anstrengung, Ausdauer) auch von Lehrkräften bei der Notengebung honoriert werden. Darüber hinaus fungieren Noten als saliente Form der Leistungsrückmeldung für Schülerinnen und Schüler, dienen daher als Quelle zur Bildung akademischer Selbstkonzepte (Marsh et al., 2005) und können vermittelt über Selbstkonzepte auch Effekte auf Interessen haben (Schurtz, Pfof, Nagengast & Artelt, 2014).

## Geschlechtsunterschiede im Selbstkonzept und Interesse

Für beide motivationalen Konstrukte wurden Mittelwertunterschiede für Mädchen und Jungen nachgewiesen. Die Selbstkonzept- und Interessenswerte in Mathematik, Physik und Chemie sind im Mittel höher bei Jungen als bei Mädchen ausgeprägt (z.B. Fredricks & Eccles, 2002; Jacobs et al., 2002; Jansen et al., 2014; Jansen et al., 2019; Schilling et al., 2006; Wilgenbusch & Merrell, 1999). Diese substanziellen Geschlechtsunterschiede im Selbstkonzept in Mathematik, Physik und Chemie sowie im Interesse in Mathematik bestehen auch unter Kontrolle korrespondierender schulischer Leistungen (z.B. Jansen et al., 2014 [Testleistung, Noten]; Parker et al., 2018 [Testleistung]). In Biologie wurden in bisherigen Studien keine Geschlechtsunterschiede im Selbstkonzept und Interesse festgestellt (Jansen et al., 2014, 2019).

In den sprachlichen Fächern geben Mädchen typischerweise höhere motivationale Werte an (z.B. Arens & Jansen, 2016; Böhme et al., 2016; Schilling et al., 2006). Unter Kontrolle von Leistungen zeigt sich für die verbale Domäne jedoch ein uneinheitliches Bild: So lassen sich Unterschiede in verbalen Selbstkonzepten zu einem substantiellen Anteil (Parker et al., 2018 [Testleistung]) oder sogar vollständig (Schilling et al., 2006 [Noten]) auf Unterschiede in den Leistungen von Mädchen und Jungen zurückführen. In einer Studie von Skaalvik und Skaalvik (2004) zeigten Jungen nach Kontrolle der Noten sogar ein höheres muttersprachliches Selbstkonzept als Mädchen; für das Interesse hingegen bestand der Vorteil der Mädchen auch nach Kontrolle der Noten als Leistungsindikator.

Zur Erklärung von Geschlechtsunterschieden im Selbstkonzept und Interesse werden insbesondere die in der Gesellschaft vorherrschenden Geschlechterstereotype als bedeutsam erachtet (z.B. Hannover & Wolter, 2019; Kessels & Heyder, 2018). Geschlechterstereotype sind verallgemeinernde Überzeugungen darüber, wie Jungen und Mädchen bzw. Männer und Frauen (angeblich) sind (*deskriptive Normen*) oder sein sollten (*präskriptive Normen*; Hannover & Wolter, 2019). Mathematisch-naturwissenschaftliche Fächer werden häufig als eher maskulin bzw. als „für Jungen“ wahrgenommen (z.B. Cvencek, Meltzoff & Greenwald, 2011; Kessels, Rau & Hannover, 2006), während sprachliche Fächer als typisch weiblich konnotiert sind (z.B. Schmenk, 2004). Zudem wird in den letzten Jahren vermehrt die These diskutiert, dass Lernen und schulische Anstrengung allgemein mit Weiblichkeit assoziiert werden (z.B. Hannover & Kessels, 2011). Solche schulbezogenen Geschlechterstereotype können über verschiedene Prozesse die schulfachbezogenen Selbstkonzepte und Interessen von Lernenden prägen. Zum einen üben Eltern, Lehrkräfte und Gleichaltrige mit ihren (geschlechtsspezifischen) Einstellungen und Verhaltenswei-

sen als Rollenmodelle einen Einfluss aus (Bandura & Walters, 1977). Zum anderen erfahren Kinder und Jugendliche durch die Reaktionen dieser Bezugspersonen auf ihr eigenes Verhalten, welche Verhaltensweisen, Attribute und Interessen als angemessen für sie eingeschätzt werden (vgl. Kessels & Heyder, 2018). So belegen verschiedene Studien beispielsweise, dass Lehrkräfte die Leistungen von Schülerinnen und Schülern unterschiedlich wahrnehmen. Die mathematischen Fähigkeiten von Jungen werden häufig über- und die von Mädchen unterschätzt (z.B. Riegle-Crumb & Humphries, 2012; Rubie-Davies & Peterson, 2016). Im sprachlichen Bereich sprechen Lehrkräfte hingegen Mädchen bessere Leistungen zu als Jungen und benoten diese auch besser – auch nach Kontrolle standardisierter Testleistungen (z.B. Lorenz, Gentrup, Kristen, Stanat & Kogan, 2016; Muntoni & Retelsdorf, 2018; Rüdiger, Jansen & Rjosk, 2021). Wenn Lernende solche geschlechtsspezifischen Differenzen in den Leistungseinschätzungen ihrer Lehrkräfte wahrnehmen, kann sich dies unter anderem auf ihre schulfachbezogenen Selbstkonzepte auswirken (Dickhäuser & Stiensmeier-Pelster, 2003; Retelsdorf, Schwartz & Asbrock, 2015).

## Selbstkonzept und Interesse im Kohortentrend

Im Rahmen der KMK-Gesamtstrategie zum Bildungsmonitoring werden die Studien des nationalen und internationalen Bildungsmonitorings in Deutschland in regelmäßigen Abständen durchgeführt (KMK, 2015). So können nicht nur Veränderungen in Kompetenzen, sondern auch Veränderungen in motivationalen Merkmalen wie den fachbezogenen Selbstkonzepten und Interessen im Zeitverlauf auf Populationsebene untersucht werden. Eine systematische Untersuchung von Kohortentrends in schulischer Motivation spezifisch für Mädchen und Jungen ist zum einen praktisch und bildungspolitisch relevant, da in den vergangenen Jahren eine Vielzahl an Förderprogrammen etabliert wurde, die insbesondere das Interesse von Mädchen und Frauen an den MINT-Fächern stärken sollen. Zum anderen sind geschlechtsspezifische Analysen von Kohortenunterschieden vor dem Hintergrund der öffentlichkeitswirksam diskutierten Frage nach Jungen als „neue Bildungsverlierer“ von Interesse (vgl. Hannover & Kessels, 2011). Diese Diskussion basiert auf der Beobachtung, dass Jungen seit einiger Zeit im Vergleich zu Mädchen schlechtere Noten in der Schule erhalten, seltener Gymnasien besuchen und die Schule häufiger ohne Abschluss verlassen (z.B. Hannover & Kessels, 2011; Helbig, 2012). Offen ist in diesem Zusammenhang die Frage, ob Jungen auch in ihrer motivationalen Entwicklung zurückfallen.

Bislang liegen nur wenige empirische Befunde zu Kohortenunterschieden in fachbezogenen Selbstkonzepten und Interessen für Sekundarschülerinnen und -schüler vor (für eine Übersicht der Studien siehe ESM 1). Überraschenderweise zeichnen die Befunde schon für kurze Zeiträume (drei [BERLIN-Studie], sechs [IQB-Bildungstrends] oder neun Jahre [PISA]) gemittelt über Mädchen und Jungen ein eher ungünstiges Bild, wonach spätere Kohorten tendenziell eine geringere schulfachbezogene Motivation in Mathematik und Deutsch zeigten als frühere Kohorten. Die wenigen Studien, die Kohortenunterschiede im Selbstkonzept und Interesse getrennt für Mädchen und Jungen untersuchten, deuten auf geschlechtsspezifische Befunde hin, die allerdings über die Studien hinweg uneinheitlich sind: Für das Fach *Mathematik* fanden sich in den PISA-Studien zwischen den Jahren 2003 und 2012 weder bei Mädchen noch bei Jungen substantielle Veränderungen im Selbstkonzept (Schiepe-Tiska & Schmidtner, 2013); im IQB-Bildungstrend wurde jedoch ein geringfügiger Anstieg des Selbstkonzepts der Mädchen und ein Rückgang des Selbstkonzepts der Jungen zwischen den Jahren 2012 und 2018 deutlich (Jansen et al., 2019). Für Freude und Interesse in Mathematik zeigte sich hingegen sowohl bei PISA als auch im IQB-Bildungstrend für beide Geschlechter ein negativer Trend mit niedrigeren Werten zum zweiten Erhebungszeitpunkt (Jansen et al., 2019; Schiepe-Tiska & Schmidtner, 2013).

In den *Naturwissenschaften* gingen in den PISA-Studien bedeutsame Rückgänge in der Selbstwirksamkeitserwartung – also der subjektiven Wahrscheinlichkeitseinschätzung, neue und/oder schwierige Situationen mit den eigenen Kompetenzen bewältigen zu können (zur Frage der konzeptionellen Abgrenzung von Selbstwirksamkeit und Selbstkonzept siehe Möller & Trautwein, 2009) – und im Interesse zwischen 2006 und 2015 auf signifikant negative Veränderungen bei den Mädchen zurück, bei Jungen hingegen blieben die Werte im Mittel stabil (Schiepe-Tiska et al., 2016). Naturwissenschaftliche Selbstwirksamkeitserwartungen und Interessen wurden in den PISA-Studien jedoch nicht fachbezogen, sondern global für alle naturwissenschaftlichen Fächer zusammen erfasst. Die Ergebnisse der IQB-Bildungstrendstudien, in denen die naturwissenschaftlichen Fächer getrennt betrachtet werden, zeigen, dass die signifikanten Veränderungen von 2012 zu 2018 überwiegend auf eine Verringerung der Selbstkonzept- und Interessenswerte der Jungen in Mathematik, Physik und Chemie zurückzuführen waren. Für Mädchen fanden sich nur wenige und nur sehr geringe Veränderungen zwischen den beiden Erhebungszeitpunkten (Jansen et al., 2019).

Auch für den *sprachlichen Bereich* weisen erste Befunde für Mädchen und Jungen gemittelt auf eine niedrigere Ausprägung im Selbstkonzept und intrinsischen Wert im Fach Deutsch in späteren Kohorten hin (2011 zu 2014; Jan-

sen et al., 2017). Die PISA-Studien ergaben zudem für die allgemeine (nicht schulspezifische) Lesefreude im Zeitraum von 2000 zu 2009 einen signifikant positiven, für den Zeitraum von 2009 zu 2018 jedoch einen signifikant negativen Trend (Artelt, Naumann & Schneider, 2010; Diedrich et al., 2019). Anschließende Analysen getrennt für Mädchen und Jungen deuten auf einen Anstieg der allgemeinen Lesefreude von 2000 zu 2009 ausschließlich bei Jungen hin, wobei von 2009 zu 2018 die allgemeine Lesefreude sowohl bei Jungen als auch bei Mädchen abgenommen hat.

## Veränderungen schulischer Leistungen und ihr möglicher Zusammenhang mit Kohortentrends in schulischer Motivation

Auch die Ergebnisse von Kohortentrends in schulischer Leistung deuten auf geschlechtsspezifische Unterschiede hin (siehe ESM 1). Ergebnisse von Kompetenztests zeigen beispielsweise in den Daten der IQB-Bildungstrendstudien einen Rückgang der Testleistungen in Mathematik zwischen 2012 und 2018 für Jungen, jedoch nicht für Mädchen (Schipolowski, Wittig, Mahler & Stanat, 2019). In den Naturwissenschaften und im Fach Deutsch fanden sich keine substantiellen Unterschiede in den Testleistungen zwischen den Erhebungszeitpunkten – weder für die Mittelwerte der Mädchen und Jungen noch im Geschlechtsunterschied (Böhme et al., 2016; Schipolowski et al., 2019). Für Noten als weiteren Schulleistungsindikator ist uns keine Studie zu Kohortenunterschieden bekannt.

Aufgrund der substantiellen Zusammenhänge von schulfachbezogenen Leistungen und korrespondierenden Selbstkonzepten/Interessen liegt auf Grundlage der bekannten Trends in den Testleistungen die Vermutung nahe, dass sich zumindest im Fach Mathematik die Geschlechtsunterschiede in den motivationalen Variablen zugunsten der Mädchen entwickeln, was sich – wie oben beschrieben – in den Trenddaten der IQB-Bildungstrendstudien für Mathematik auch tatsächlich zeigt. Veränderungen in den Leistungen zwischen Kohorten für Mädchen und Jungen könnten somit zumindest teilweise für die gefundenen geschlechtsspezifischen Kohortenunterschiede in den motivationalen Variablen verantwortlich sein. Empirische Evidenz zu dieser Annahme existiert unseres Wissens bisher nicht.

## Forschungsfragen

Ziel der vorliegenden Studie ist es, geschlechtsspezifische Kohortentrends im Selbstkonzept in Mathematik, den na-

turwissenschaftlichen Fächern und im Fach Deutsch sowie im Interesse in Mathematik und den naturwissenschaftlichen Fächern differenziert zu untersuchen. In der vorliegenden Studie soll deshalb getrennt für die Fächer geprüft werden, ob (a) Kohortenunterschiede in den fachbezogenen Selbstkonzepten und Interessen geschlechterdifferenziell ausfallen (für die MINT-Fächer sind diese Befunde bereits bei Jansen et al., 2019 zu finden) und insbesondere, (b) ob diese auf Unterschiede in schulischen Leistungen zwischen Kohorten zurückgeführt werden können. Als Schulleistungsindikatoren werden dabei sowohl Leistungen in Kompetenztests als auch Noten berücksichtigt. Aufgrund der höheren Zusammenhänge zwischen Noten und motivationalen Variablen im Vergleich zu Testleistungen (vgl. z.B. Jansen et al., 2014; Lotz, Schneider & Sparfeldt, 2018) vermuteten wir, dass sich Kohortentrends in motivationalen Variablen stärker auf Notenunterschiede zwischen Kohorten zurückführen lassen als auf Unterschiede in den Testleistungen.

## Methode

### Stichprobe

Die Analysen basieren für die Fächer Mathematik, Physik, Chemie und Biologie auf den Daten des IQB-Ländervergleichs 2012 (Pant et al., 2015) und des IQB-Bildungstrends 2018 (Stanat, Schipolowski, Mahler, Weirich & Henschel, 2019) sowie für das Fach Deutsch auf den Daten des IQB-Ländervergleichs 2008/2009 (Köller, Knigge & Tesch, 2011) und des IQB-Bildungstrends 2015 (Stanat et al., 2018). In diesen querschnittlich angelegten Studien wurden repräsentative Daten bei Schülerinnen und Schülern am Ende der neunten Jahrgangsstufe an allgemeinbildenden Schularten sowie Förderschulen erhoben (für Details zur Stichprobe siehe ESM 2).

### Messinstrumente

**Geschlecht.** Das Geschlecht der Schülerinnen und Schüler wurde in der Schülerdemografieliste erfasst, welche durch Ansprechpersonen an der Schule ausgefüllt wurde.

**Selbstkonzept und Interesse.** Das Selbstkonzept sowie das Interesse der Schülerinnen und Schüler in Mathematik, Physik, Chemie und Biologie wurden über die Fä-

cher hinweg sowie zu den beiden Erhebungszeitpunkten (2012 und 2018) mit den gleichen vier Items erhoben (z.B. Selbstkonzept: „In [Fach] lerne ich schnell“; Interesse: „Für [Fach] interessiere ich mich“). Für das Selbstkonzept in Deutsch wurden sieben Items in die vorliegende Studie einbezogen, die zu beiden Erhebungszeitpunkten (2009 und 2015) erfasst wurden.<sup>1</sup> Auf einer vierstufigen Skala beurteilten die Jugendlichen jeweils, inwieweit die Aussagen auf sie selbst zutreffen (1 [*trifft überhaupt nicht zu*] bis 4 [*trifft völlig zu*]). Die internen Konsistenzen für die einzelnen Skalen sind mindestens als akzeptabel, überwiegend aber als sehr gut zu bewerten (ESM 2). Für die nachfolgenden Analysen wurde aus den schulfachbezogenen Selbstkonzept- bzw. Interessensitems pro Erhebungszeitpunkt für jede Schülerin und jeden Schüler der Skalenmittelwert gebildet.

**Leistungen.** In den IQB-Studien wurden die mathematischen, naturwissenschaftlichen sowie sprachlichen Kompetenzen der Schülerinnen und Schüler mit Testaufgaben auf Grundlage der entsprechenden Bildungsstandards der Kultusministerkonferenz erfasst. Für unsere Analysen verwendeten wir die Globalskala mathematischer Kompetenz, in den naturwissenschaftlichen Fächern die Kompetenzskalen für das *Fachwissen* (vgl. Mahler, Schipolowski & Weirich, 2019). Für die Analysen im Fach Deutsch wurden – vergleichbar mit dem Fach Mathematik – Testwerte der Schülerinnen und Schüler auf einer Globalskala verwendet.<sup>2</sup> Die Kompetenztestwerte werden auf der KMK-Bildungsstandardmetrik berichtet (für Gesamtdeutschland zum jeweils ersten Erhebungszeitpunkt des Trends:  $M = 500$ ,  $SD = 100$ ).

Zudem wurden Ansprechpersonen an den Schulen gebeten, die Halbjahreszeugnisnoten der Schülerinnen und Schüler in den entsprechenden Fächern anzugeben. Aufgrund geringer Fallzahlen bei sehr schlechten Noten wurden *ungenügend* und *mangelhaft* für die Analysen zusammengefasst. Die schulfachbezogenen Noten wurden umgepolt, sodass höhere numerische Werte besseren Leistungen entsprechen (1 [*ungenügend* oder *mangelhaft*] bis 5 [*sehr gut*]).

### Datenanalyse

**Schätzung der Kompetenztestwerte und multiple Imputation fehlender Werte.** Zur Schätzung der Kompetenztestwerte wurde in den IQB-Studien der auf dem Prinzip der multiplen Imputation beruhende *Plausible Values*

<sup>1</sup> Im Jahr 2009 wurde das Deutschselbstkonzept mit elf Items, im Jahr 2015 mit sieben Items erfasst. In diesem Beitrag wurden nur die sieben Items in die Analysen einbezogen, die zu beiden Zeitpunkten identisch waren.

<sup>2</sup> In den Berichtsbänden zum IQB-Ländervergleich 2008/2009 und zum IQB-Bildungstrend 2015 werden Ergebnisse getrennt für die drei Kompetenzbereiche Lesen, Zuhören und Orthografie berichtet. Die hier verwendete Globalskala wurde für den vorliegenden Beitrag gebildet.

(PV) Ansatz genutzt (für eine Beschreibung des genauen Vorgehens vgl. Becker, Weirich, Mahler & Sachse, 2019; Sachse, Haag & Weirich, 2016). In allen Datensätzen wurden 15 PVs pro Fach gezogen mit Ausnahme des IQB-Ländervergleichs 2009, bei dem nur 5 PVs gezogen wurden.<sup>3</sup>

Zur Imputation fehlender Werte für alle anderen Variablen (Geschlecht, Skalenmittelwerte für schulfachbezogene Selbstkonzepte und Interessen, Noten) wurde das Verfahren *Multivariate Imputation by Chained Equations* verwendet, das im R-Paket *mice* implementiert ist (van Buuren & Groothuis-Oudshoorn, 2011). Dabei wurden die fehlenden Werte einer Variablen anhand der vorhandenen Informationen zu allen anderen Variablen des Gesamtdatensatzes des jeweiligen IQB-Ländervergleichs bzw. IQB-Bildungstrends geschätzt (Becker et al., 2019).

**Statistische Analysemodelle.** Deskriptive Analysen wurden mit dem R-Paket *eatAnalysis* (Becker, Weirich & Sachse, 2021) vorgenommen, trendbezogene Mittelwerts- und Regressionsanalysen mit dem Paket *eatRep* (Weirich, Hecht & Becker, 2021). Die Pakete sind für die Analyse großer Schulleistungstudien unter Berücksichtigung ihrer spezifischen Datenstruktur konzipiert. So wurden bei den Berechnungen Stichprobengewichte genutzt und durch Anwendung des *Jackknife*-Verfahrens (vgl. Weirich, Hecht, Becker & Zitzmann, 2020) der Standardfehler zur Berücksichtigung der Mehrebenenstruktur der Daten bzw. der komplexen Stichprobenziehung korrigiert. Außerdem wurden mit dem Paket die trendbezogenen Mittelwerts- und Regressionsanalysen separat für die imputierten Datensätze<sup>4</sup> durchgeführt und die Koeffizienten anschließend gepoolt (Mittelwerte und Regressionskoeffizienten anhand der Regeln von Rubin, 2003;  $R^2$  entsprechend Harel, 2009).

In einem ersten Analyseschritt wurden Mittelwerte, Standardfehler und Standardabweichungen der fachbezogenen Selbstkonzepte und Interessen, Testleistungen (PVs) und umgepolten Noten getrennt für Mädchen und Jungen sowie für die Erhebungszeitpunkte bestimmt. Zusätzlich wurden die Differenzen der Mittelwerte zwischen den Erhebungszeitpunkten mit dazugehörigen Standardfehlern berechnet und die Effektstärke Cohens  $d$  (Cohen,

1988) sowie die Signifikanz der Veränderungen ermittelt.<sup>5</sup> Die Einordnung der Stärke von Effekten basiert auf den Empfehlungen von Gignac & Szodorai (2016) mit  $r = .10$  (entspricht  $d = 0.20$ ) als kleinem,  $r = .20$  ( $d = 0.40$ ) als mittlerem und  $r = .30$  ( $d = 0.60$ ) als großem Effekt. Dementsprechend betrachten wir statistisch signifikante Veränderungen mit einer Stärke ab  $d = 0.20$  als bedeutsam.

Um zu untersuchen, ob die geschlechtsspezifischen Kohortentrends in den fachbezogenen Selbstkonzepten und Interessen auch nach Kontrolle von Leistungen bestehen bleiben, wurden in einem zweiten Schritt für jede abhängige Variable (Selbstkonzept/Interesse in den einzelnen Fächern) separate Regressionsanalysen durchgeführt. Im Modell 0 (M0) wurden zunächst die geschlechtsspezifischen Kohortentrends ohne Kontrolle von Leistungsmaßen bestimmt. Die in diesen Modellen geschätzten Parameter des Erhebungszeitpunkts (entspricht dem Kohortentrend für Jungen) und der Interaktion aus Geschlecht und Erhebungszeitpunkt (entspricht dem Unterschied des Kohortentrends zwischen Mädchen und Jungen, also dem geschlechtsspezifischen Kohortentrend) dienten als Vergleichsmaß für die Modelle 1 (M1), 2 (M2) und 3 (M3), in denen schrittweise verschiedene Leistungsmaße kontrolliert wurden. In M1 berücksichtigten wir die schulfachbezogenen Testleistungen (PVs), in M2 die schulfachbezogenen, umgepolten Noten. In M3 bezogen wir beide Leistungsmaße simultan ein. In M1 bis M3 wurden neben den Haupteffekten der Leistungsmaße auch Leistungsunterschiede zwischen den Kohorten durch den Einbezug von Interaktionstermen (Leistungsmaß x Erhebungszeitpunkt) berücksichtigt. Die Veränderung der geschätzten Parameter für Erhebungszeitpunkt und Geschlecht\*Erhebungszeitpunkt gegenüber M0 wird in der Ergebnisdarstellung mit  $\Delta b$  bezeichnet.

Durch die Dummy-Kodierung der Variablen Geschlecht (0 = Junge, 1 = Mädchen) und Erhebungszeitpunkt (0 = erster Erhebungszeitpunkt, 1 = zweiter Erhebungszeitpunkt) bilden Jungen zum ersten Erhebungszeitpunkt in allen Modellen die Referenzgruppe. Alle metrischen Prädiktoren (PVs und umgepolte Noten) gingen z-standardisiert in die Analysen ein.

<sup>3</sup> Technische Voraussetzung für Trendanalysen mit *Plausible Values* ist, dass für beide Erhebungszeitpunkte, die verglichen werden, die gleiche Anzahl *Plausible Values* genutzt wird. In der vorliegenden Studie basieren die Testleistungen in den Analysen für das Fach Deutsch daher sowohl für das Jahr 2009 als auch das Jahr 2015 auf je 5 *Plausible Values*. Im Jahr 2015 wurden für das Fach Deutsch also nur die ersten 5 *Plausible Values* einbezogen, obwohl zunächst 15 *Plausible Values* gezogen worden waren.

<sup>4</sup> Da sowohl *Plausible Values* als auch imputierte Fragebogendaten verwendet wurden, wurden diese so kombiniert, dass jeweils die *n*te Imputation der Fragebogendaten mit der *n*ten *Plausible Value* Ziehung der Leistungsdaten zusammengeführt wurde (also 1. Imputation mit 1. PV, 2. Imputation mit 2. PV usw.).

<sup>5</sup> Zum Teil sind diese Befunde bereits äquivalent in den Berichtsbänden der IQB-Bildungstrendstudien berichtet worden (Böhme et al., 2016; Jansen et al., 2019; Schipolowski et al., 2019). Für alle Fächer und Variablen ergeben sich leichte Abweichungen zu den Werten in den Berichtsbänden durch die Verwendung einer imputierten Geschlechtervariable sowie die Verwendung schulfachspezifischer Gewichte für Selbstkonzepte und Interessen. Im Fach Deutsch wurde zusätzlich die Selbstkonzeptvariable neu gebildet (siehe Fußnote 1) und ein Globalmodell für die schulischen Kompetenzen im Fach Deutsch spezifiziert (siehe Fußnote 2).

## Ergebnisse

Im ESM 2 finden sich die Korrelationen der zentralen kontinuierlichen Variablen (Selbstkonzept, Interesse, Testleistung, Note) in den betrachteten Fächern separat für Mädchen und Jungen und die beiden Erhebungszeitpunkte. ESM 3 enthält Abbildungen, in denen die Selbstkonzept- und Interessenswerte für Mädchen und Jungen je Note beziehungsweise Kompetenzstufe nach Erhebungszeitpunkten dargestellt sind. Im Folgenden werden die Ergebnisse der Mittelwertvergleiche und Regressionsanalysen nacheinander für Mathematik, die naturwissenschaftlichen Fächer und Deutsch vorgestellt. Für eine bessere Verständlichkeit der Ergebnisse werden diese für das Selbstkonzept im Fach Mathematik ausführlicher beschrieben.

### Mathematik

Die Mittelwertvergleiche (ESM 2) zwischen den Kohorten zeigen, dass das Selbstkonzept bei Mädchen von T1 zu T2 geringfügig angestiegen ( $d = 0.06$ ), bei Jungen hingegen zurückgegangen ist ( $d = -0.16$ ). Für das Interesse ergab sich ein ähnliches Befundmuster, wobei der negative Trend für Jungen ( $d = -0.31$ ) stärker ausfiel als beim Selbstkonzept. In den Leistungsmaßen zeigte sich ebenfalls eine geschlechtsspezifische Entwicklung im Trend: Während die Testleistungen der Mädchen zu beiden Erhebungszeitpunkten vergleichbar waren ( $d = 0.03$ ), erzielten Jungen 2018 geringfügig schlechtere Testleistungen als noch 2012 ( $d = -0.06$ ). Bei den Noten hingegen zeigte sich im Trend eine geringfügige Verbesserung für Mädchen ( $d = 0.08$ ), jedoch keine Veränderung für Jungen ( $d = -0.01$ ).

Die Regressionsanalysen (Tab. 1) geben Aufschluss über die Kohortentrends im Selbstkonzept und Interesse unter Kontrolle der Leistungsmaße. M0 im linken Teil der Tabelle mit den Prädiktoren Geschlecht, Erhebungszeitpunkt sowie deren Interaktion stellt die geschlechtsspezifischen Mittelwertvergleiche für das Mathematikselbstkonzept regressionsanalytisch dar: Das Intercept des Modells ( $b = 2.85$ ) entspricht dem Selbstkonzeptwert der Referenzgruppe, also Jungen zu T1 ( $M = 2.85$ ). Der Effekt des Geschlechts ( $b = -0.44$ ) bildet somit den Unterschied zwischen Mädchen und Jungen zu T1 ab, womit sich in Kombination mit dem Intercept der Mittelwert der Mädchen für T1 berechnen lässt ( $2.85 - 0.44 = 2.41$ ). Der Effekt des Erhebungszeitpunkts (EZ;  $b = -0.13$ ) gibt den Mittelwertunterschied zwischen T1 und T2 für Jungen an ( $\Delta M = -0.13$ ). Der Interaktionseffekt aus Geschlecht und Erhebungszeitpunkt ( $b = 0.19$ ) schließlich beziffert den Unterschied zwischen Jungen und Mädchen im Kohortentrend, woraus sich auch der Kohortentrend für Mädchen ergibt ( $-0.13 + 0.19 = 0.06$ ). Insgesamt wurden durch das Geschlecht und den Erhe-

bungszeitpunkt etwa 5% der Varianz im mathematischen Selbstkonzept erklärt.

In M1 wurde zusätzlich die Testleistung als Ausgangswert (PV) und Trend (EZ\*PV) aufgenommen. Es zeigte sich ein positiver Effekt der Testleistung auf das Selbstkonzept ( $b = 0.26$ ). Die Aufnahme der Testleistung führte aber nicht zu einer Veränderung des Kohortentrends für Jungen ( $\Delta b = -0.01$ , EZ im Vergleich zu M0) oder des geschlechtsspezifischen Kohortentrends ( $\Delta b = 0.02$ , Geschlecht\*EZ im Vergleich zu M0). Die Varianzaufklärung von M1 lag bei 16%. M2 mit der Mathematiknote als einziger Leistungsindikator erklärte einen wesentlich höheren Anteil der Varianz im mathematischen Selbstkonzept ( $R^2 = .40$ ), wirkte sich aber ebenfalls nicht auf das Muster des Kohortentrends für Jungen ( $\Delta b = 0.00$ ) oder den geschlechtsspezifischen Kohortentrend aus ( $\Delta b = 0.05$ ). Auch nach Kontrolle beider Leistungsmaße in M3 zeigte sich weiterhin ein negativer Kohortentrend für Jungen ( $b = -0.13$ ) und eine signifikante Interaktion zwischen Erhebungszeitpunkt und Geschlecht ( $b = 0.15$ ), die auf differenzielle Kohortentrends für Mädchen und Jungen hinweist. Beide Koeffizienten haben sich nach Kontrolle der beiden Leistungsmaße nicht bedeutsam verändert (EZ:  $\Delta b = 0.00$ ; Geschlecht\*EZ:  $\Delta b = 0.04$ ).

Die Befunde der Regressionsanalysen für das Interesse an Mathematik (rechter Teil Tab. 1) zeigen ein ähnliches Befundmuster. Auch hier blieb zudem nach Kontrolle beider Leistungsmaße ein negativer Kohortentrend für Jungen ( $b = -0.26$ ,  $\Delta b = 0.00$ ) und ein geschlechtsspezifischer Kohortentrend ( $b = 0.19$ ,  $\Delta b = 0.03$ ) bestehen.

### Naturwissenschaften

**Physik.** Im Fach Physik ergaben die Mittelwertvergleiche (ESM 2) keine bzw. eine geringfügig positive Veränderung des Selbstkonzepts und Interesses bei Mädchen, während das Selbstkonzept und Interesse bei Jungen von T1 zu T2 abnahm. Die Testleistungen unterschieden sich für beide Geschlechter nicht zwischen den Kohorten. Ebenso blieben die Noten der Jungen vergleichbar. Die Noten der Mädchen verbesserten sich hingegen von T1 zu T2 geringfügig. Den Regressionsanalysen (Tab. 1) zufolge veränderte sich für das Selbstkonzept und für das Interesse weder der Kohortentrend für Jungen noch der Geschlechtsunterschied im Kohortentrend nach Kontrolle der Leistungsmaße.

**Chemie.** Auch für das Fach Chemie zeigten die Mittelwertvergleiche (ESM 2) keine bedeutsame Veränderung von Selbstkonzept und Interesse bei Mädchen zwischen den Kohorten, wohingegen das Selbstkonzept und das Interesse bei Jungen leicht zurückgingen. Bei den Testleistungen waren für beide Geschlechter leichte Rückgänge zu verzeichnen. Gleichzeitig erhielten Mädchen im Fach

**Tabelle 1.** Ergebnisse der Regressionsanalysen ohne Kontrolle von Leistung (M0), unter Kontrolle der Testleistung (M1), der Noten (M2) und beider Leistungsmaße (M3) – separat für die Selbstkonzepte und Interessen in Mathematik, Physik, Chemie, Biologie und Deutsch

Variable	Selbstkonzept						Interesse									
	M0		M1		M2		M3		M0		M1		M2		M3	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
<b>Mathematik</b>																
Intercept	2.85*	(0.02)	2.81*	(0.02)	2.86*	(0.02)	2.85*	(0.02)	2.66*	(0.04)	2.64*	(0.03)	2.67*	(0.03)	2.67*	(0.03)
Geschlecht	-0.44*	(0.02)	-0.40*	(0.02)	-0.45*	(0.02)	-0.43*	(0.02)	-0.37*	(0.03)	-0.35*	(0.03)	-0.38*	(0.03)	-0.38*	(0.03)
EZ	-0.13*	(0.03)	-0.12*	(0.03)	-0.13*	(0.02)	-0.13*	(0.02)	-0.26*	(0.07)	-0.25*	(0.04)	-0.26*	(0.04)	-0.26*	(0.04)
PV			0.26*	(0.01)			0.08*	(0.02)			0.15*	(0.02)			0.02	(0.02)
Note					0.51*	(0.01)	0.48*	(0.01)					0.35*	(0.01)	0.34*	(0.02)
Geschlecht*EZ	0.19*	(0.03)	0.17*	(0.03)	0.14*	(0.02)	0.15*	(0.03)	0.22*	(0.04)	0.20*	(0.04)	0.19*	(0.03)	0.19*	(0.03)
EZ*PV			0.05*	(0.02)			0.05*	(0.02)			-0.01	(0.02)			-0.01	(0.02)
EZ*Note					-0.01	(0.02)	-0.03	(0.02)					-0.03*	(0.02)	-0.03	(0.02)
R <sup>2</sup>	0.05		0.16		0.40		0.42		0.04		0.07		0.20		0.20	
<b>Physik</b>																
Intercept	2.75*	(0.03)	2.74*	(0.02)	2.74*	(0.02)	2.74*	(0.02)	2.48*	(0.02)	2.48*	(0.02)	2.48*	(0.02)	2.47*	(0.02)
Geschlecht	-0.44*	(0.02)	-0.44*	(0.02)	-0.43*	(0.02)	-0.43*	(0.02)	-0.53*	(0.02)	-0.53*	(0.02)	-0.53*	(0.02)	-0.53*	(0.02)
EZ	-0.11*	(0.04)	-0.11*	(0.02)	-0.10*	(0.02)	-0.10*	(0.02)	-0.11*	(0.03)	-0.10*	(0.02)	-0.10*	(0.02)	-0.10*	(0.02)
PV			0.13*	(0.02)			0.03*	(0.02)			0.12*	(0.02)			0.05*	(0.02)
Note					0.37*	(0.01)	0.36*	(0.01)					0.27*	(0.01)	0.26*	(0.01)
Geschlecht*EZ	0.15*	(0.02)	0.14*	(0.02)	0.10*	(0.02)	0.10*	(0.02)	0.16*	(0.03)	0.16*	(0.03)	0.12*	(0.03)	0.13*	(0.03)
EZ*PV			0.05	(0.04)			0.03	(0.03)			0.01	(0.04)			-0.01	(0.03)
EZ*Note					-0.02	(0.02)	-0.04*	(0.01)					-0.03	(0.02)	-0.03	(0.02)
R <sup>2</sup>	0.06		0.10		0.28		0.28		0.08		0.10		0.17		0.18	
<b>Chemie</b>																
Intercept	2.66*	(0.03)	2.66*	(0.02)	2.69*	(0.01)	2.69*	(0.01)	2.39*	(0.03)	2.39*	(0.02)	2.41*	(0.02)	2.41*	(0.02)
Geschlecht	-0.18*	(0.03)	-0.19*	(0.03)	-0.20*	(0.02)	-0.20*	(0.02)	-0.22*	(0.03)	-0.23*	(0.03)	-0.24*	(0.02)	-0.24*	(0.02)
EZ	-0.08*	(0.04)	-0.07*	(0.02)	-0.08*	(0.02)	-0.08*	(0.02)	-0.07*	(0.03)	-0.06*	(0.02)	-0.07*	(0.02)	-0.07*	(0.02)
PV			0.14*	(0.01)			0.01	(0.01)			0.14*	(0.02)			0.05*	(0.02)
Note					0.42*	(0.01)	0.41*	(0.01)					0.31*	(0.02)	0.30*	(0.02)
Geschlecht*EZ	0.06	(0.03)	0.05	(0.02)	0.01	(0.03)	0.01	(0.03)	0.08*	(0.03)	0.08*	(0.03)	0.05	(0.03)	0.05	(0.03)
EZ*PV			0.05	(0.02)			0.04*	(0.02)			0.02	(0.05)			0.00	(0.04)



**Tabelle 1.** Ergebnisse der Regressionsanalysen ohne Kontrolle von Leistung (M0), unter Kontrolle der Testleistung (M1), der Noten (M2) und beider Leistungsmaße (M3) – separat für die Selbstkonzepte und Interessen in Mathematik, Physik, Chemie, Biologie und Deutsch (Fortsetzung)

Variable	Selbstkonzept						Interesse											
	M0		M1		M2		M3		M0		M1		M2		M3			
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)		
EZ*Note	0.01		0.06		0.28		-0.04*	(0.02)	0.01		0.04		-0.02	(0.03)	0.14		-0.02	(0.02)
R <sup>2</sup>							0.38		0.01		0.04		0.14		0.15			
<b>Biologie</b>																		
Intercept	2.81*	(0.02)	2.82*	(0.01)	2.82*	(0.01)	2.83*	(0.01)	2.48*	(0.02)	2.49*	(0.02)	2.49*	(0.02)	2.49*	(0.02)	2.49*	(0.02)
Geschlecht	0.02	(0.02)	-0.01	(0.02)	-0.02	(0.02)	-0.03	(0.02)	0.05*	(0.02)	0.02	(0.02)	0.02	(0.02)	0.02	(0.02)	0.01	(0.02)
EZ	-0.09*	(0.04)	-0.08*	(0.02)	-0.08*	(0.02)	-0.08*	(0.02)	-0.10*	(0.03)	-0.10*	(0.02)	-0.09*	(0.03)	-0.09*	(0.02)	-0.09*	(0.02)
PV			0.13*	(0.01)			0.06*	(0.01)			0.12*	(0.02)			0.07*	(0.01)		
Note					0.30*	(0.01)	0.28*	(0.01)			0.21*	(0.02)			0.19*	(0.02)		
Geschlecht*EZ	0.07*	(0.03)	0.06*	(0.03)	0.04	(0.03)	0.04	(0.03)	0.15*	(0.03)	0.14*	(0.03)	0.13*	(0.03)	0.13*	(0.03)	0.13*	(0.03)
EZ*PV			0.05*	(0.02)			0.03	(0.02)			0.02	(0.03)			0.00	(0.02)		
EZ*Note					0.00	(0.01)	-0.01	(0.01)					-0.01	(0.03)			-0.01	(0.02)
R <sup>2</sup>			0.01		0.06		0.20		0.01		0.04		0.08		0.09			
<b>Deutsch</b>																		
Intercept	3.14*	(0.01)	3.16*	(0.01)	3.24*	(0.01)	3.24*	(0.01)										
Geschlecht	0.18*	(0.01)	0.13*	(0.01)	0.05*	(0.01)	0.04*	(0.01)										
EZ	0.04*	(0.01)	0.05*	(0.01)	0.00	(0.01)	0.01	(0.01)										
PV			0.15*	(0.01)			0.06*	(0.01)										
Note					0.28*	(0.01)	0.26*	(0.01)										
Geschlecht*EZ	0.00	(0.02)	0.00	(0.02)	0.01	(0.01)	0.01	(0.01)										
EZ*PV			0.00	(0.01)			0.01	(0.01)										
EZ*Note					-0.04*	(0.01)	-0.04*	(0.01)										
R <sup>2</sup>			0.03		0.09		0.21		0.03		0.09		0.22					

Anmerkungen: \*p < .05; EZ = Erhebungszeitpunkt; PV = Testleistung; Die Ergebnisse zusätzlicher Regressionsanalysen unter Berücksichtigung der Dreifach-Interaktionen aus Geschlecht\*EZ\*PV bzw. Geschlecht\*EZ\*Noten sind im ESM 4 aufbereitet.

Chemie zu T2 etwas bessere Noten als zu T1. Die Noten der Jungen blieben konstant. Die Regressionsanalysen (Tab. 1) für das Fach Chemie ergaben ein etwas anderes Ergebnismuster als im Fach Physik. Zwar veränderte sich der Kohortentrend für Jungen für beide motivationalen Merkmale auch im Fach Chemie nach Kontrolle der Leistungsmaße nicht bedeutsam. Die Unterschiede in den Kohortentrends für Selbstkonzept und Interesse zwischen Mädchen und Jungen reduzierte sich im Fach Chemie jedoch minimal und waren in M2 und M3 statistisch nicht mehr signifikant. Der Kohortentrend im Selbstkonzept unterschied sich allerdings bereits in M0 nicht signifikant zwischen Mädchen und Jungen.

**Biologie.** Wie bereits in den anderen naturwissenschaftlichen Fächern ergaben die Mittelwertvergleiche (ESM 2) auch für das Fach Biologie keine bzw. geringfügig positive Veränderungen des Selbstkonzepts und Interesses bei Mädchen, aber leichte Rückgänge zwischen den Kohorten für Jungen. Die Testleistungen und Noten sind für beide Geschlechter zwischen den Kohorten konstant geblieben. Die Regressionsanalysen (Tab. 1) ergaben für das Fach Biologie ein ähnliches Bild wie im Fach Chemie. Auch hier blieben die Kohortentrends für Jungen nach Kontrolle der Leistungsmaße stabil. Die Unterschiede in den Kohortentrends zwischen Mädchen und Jungen verringerten sich sowohl für das Selbstkonzept als auch das Interesse minimal nach Kontrolle der Leistungsmaße und waren für das Selbstkonzept in M2 und M3 nicht mehr statistisch signifikant.

## Deutsch

Im Fach Deutsch ergaben die Mittelwertvergleiche (ESM 2) einen geringfügigen Anstieg in den Selbstkonzepten für beide Geschlechter. Während sich in den Testleistungen für beide Geschlechter keine Veränderungen zeigten, fand sich in Bezug auf die Noten sowohl für Mädchen als auch für Jungen ein positiver Kohortentrend mit im Mittel besseren Noten zu T2 als zu T1. Die Regressionsanalysen (Tab. 1) ergaben, dass sich der geringfügig positive Trend für Jungen nach Kontrolle der Leistungsmaße nur marginal veränderte, in M2 und M3 aber nicht mehr statistisch signifikant war. Der Interaktionseffekt aus Geschlecht und Erhebungszeitpunkt blieb unverändert und nicht signifikant.

## Robustheitschecks

Um die Robustheit der vorgestellten Analysen zu überprüfen, haben wir eine Reihe zusätzlicher Regressionsanalysen durchgeführt. Auf Basis dieser Robustheitschecks (ESM 4) kommen wir gleichbleibend zu dem Schluss, dass die teilweise beobachteten Kohortentrends im Selbstkonzept

und Interesse nicht bedeutsam auf Veränderungen in den Leistungen der Schülerinnen und Schüler zurückgeführt werden können.

## Diskussion

Ausgangspunkt unseres Beitrags waren die für uns unerwarteten Befunde, die schon für kurze Zeiträume von drei bis neun Jahren Kohortenunterschiede in schulischer Motivation berichteten. Wir untersuchten in diesem Beitrag, ob geschlechtsspezifische Kohortenunterschiede in fachbezogenen Selbstkonzepten und Interessen auftreten und insbesondere, ob sie auf Unterschiede in schulischen Leistungen zwischen diesen Kohorten zurückgeführt werden können. Die Abnahme der fachbezogenen Selbstkonzepte und Interessen der Jungen von 2012 zu 2018 in den mathematisch-naturwissenschaftlichen Fächern ließ sich nicht auf Veränderungen ihrer schulischen Leistungen zurückführen. Ebenso hing die geringfügige Zunahme des Deutschselbstkonzepts für Mädchen und Jungen nicht bedeutsam mit ihren etwas besseren Noten in 2015 im Vergleich zu 2009 zusammen.

## Kohortentrends in fachbezogenen Selbstkonzepten und Interessen

Für die Selbstkonzepte in Mathematik und in den naturwissenschaftlichen Fächern zeigten sich in unserer Studie geschlechtsspezifische Kohortentrends mit vergleichbaren oder sogar geringfügig höher ausgeprägten Selbstkonzepten und Interessen in 2018 im Vergleich zu 2012 für Mädchen, für Jungen hingegen war für alle Fächer ein Rückgang in den beiden motivationalen Variablen zu beobachten. Diese Befunde stehen nicht im Widerspruch zu den Ergebnissen der PISA-Studien, welche zeitlich vor den IQB-Bildungstrends durchgeführt wurden und zwischen den Jahren 2003 und 2012 weder bei Mädchen noch bei Jungen substantielle Veränderungen im Selbstkonzept, für das Interesse in Mathematik jedoch einen für beide Geschlechter vergleichbaren negativen Trend fanden (Schiepe-Tiska & Schmidtner, 2013). Für die naturwissenschaftlichen Fächer können die Trendergebnisse der IQB-Bildungstrend- und PISA-Studien u. a. aufgrund der fachübergreifenden Erfassung der motivationalen Konstrukte bei PISA nicht direkt miteinander verglichen werden. Allein auf Grundlage der Ergebnisse im mathematisch-naturwissenschaftlichen Bereich könnte die Annahme, dass Jungen die „neuen Bildungsverlierer“ sind, auf die schulische Motivation, spezifisch auf Selbstkonzept und Interesse, erweitert werden. Gegen eine solche globale Deutung spricht jedoch der zwar

geringfügige, aber eben doch positive Trend im Selbstkonzept im Fach Deutsch bei den Jungen. Vielmehr scheint der Rückgang der Motivation bei den Jungen auf den mathematisch-naturwissenschaftlichen Bereich beschränkt zu sein. Zur Absicherung dieser Annahme sollten in zukünftiger Forschung Trends in Selbstkonzepten und Interessen in weiteren Fächern des sprachlichen (und gegebenenfalls des musisch-künstlerischen oder gesellschaftlichen) Bereichs einbezogen werden.

Obwohl die Ausprägung der fachbezogenen Selbstkonzepte und Interessen substanziell mit den fachbezogenen schulischen Leistungen zusammenhing, konnten die gefundenen Kohortentrends in beiden motivationalen Variablen nicht bedeutsam auf Leistungsveränderungen zwischen den Kohorten zurückgeführt werden. Andere, in dieser Studie nicht untersuchte potentielle Veränderungen müssten somit zum besseren Verständnis geschlechtsspezifischer Trends in schulischer Motivation herangezogen werden. Während beispielsweise Veränderungen in der Qualität des Bildungssystems (z.B. der Unterrichtsqualität), der Klassenzusammensetzung (z.B. aufgrund von Schulstrukturreformen) oder der gesellschaftlichen Wertschätzung gegenüber Schule und einzelnen Fächern Schülerinnen und Schüler gleichermaßen betreffen und somit für die Veränderungen der Selbstkonzepte und Interessen von Mädchen und Jungen gleichermaßen relevant sein dürften, sind auch Veränderungen denkbar, die zu geschlechtsspezifischen Trends in Selbstkonzepten und Interessen beitragen könnten.

Eine Annahme, warum sich die Rückgänge im Selbstkonzept und Interesse insbesondere in der als männlich konnotierten MINT-Domäne (und hierbei besonders im Mathematikinteresse) zeigen, könnten Veränderungen in geschlechtsspezifischen Überzeugungen sein (vgl. Gentrup & Stanat, 2021). In Studien wurde beobachtet, dass Einstellungen zu Geschlechterrollen zunehmend egalitärer geworden sind – etwa bezogen auf familiäre Care-Arbeit (Zoch & Schober, 2018). Möglicherweise weichen vermehrt auch solche schulbezogenen Geschlechterstereotype wie das Stereotyp „MINT ist für Jungen“ auf. Da Interessen eng mit der Entwicklung der (Geschlechts-)Identität von Jugendlichen verknüpft sind (*Interessen als Identitätsregulation Modell*; Kessels & Hannover, 2007), könnte eine Aufweichung des MINT-Jungen-Stereotyps dazu beitragen, dass Lernende ihre Interessen individueller und insbesondere weniger durch ihr Geschlecht beeinflusst ausbilden. Jungen könnten also weniger das Gefühl haben, gute Leistungen in den MINT-Fächern erzielen und an den MINT-Fächern interessiert sein zu müssen, um als männlich wahrgenommen zu werden. Ebenso könnte eine Identifikation mit MINT-Fächern für die Selbstdarstellung der Jungen als männlich weniger wichtig geworden sein. Eine verstärkende Wirkung könnten in diesem Zusam-

menhang die vielfältigen MINT-Programme speziell für Mädchen und Frauen haben, die das Ziel verfolgen, das Potential von Frauen für naturwissenschaftlich-technische Berufe auszuschöpfen und das Bild der MINT-Berufe in der Gesellschaft zu verändern – hin zu einer weniger männlichen Domäne.

Insgesamt ist es schwierig, die Effekte verschiedener, zeitlich parallel laufender Veränderungen (z.B. Veränderung von Geschlechterstereotypen, Implementation von geschlechtsspezifischen Förderprogrammen) zu trennen und ihre Einflüsse auf fachbezogene motivationale Variablen zu isolieren. Um die negativen Motivationstrends besser zu verstehen und ihnen entgegenwirken zu können, ist es dennoch sinnvoll, potentiell relevante Konstrukte und deren Veränderungen zu erfassen sowie diese Veränderungen mit (auch längerfristigen) Trends in motivationalen Variablen in Verbindung zu setzen.

## Implikationen für Forschung und Praxis

Aus den Befunden unserer Studie ergeben sich einige Implikationen für die zukünftige Forschung. So ist der minimale Anstieg des Selbstkonzepts im Fach Deutsch zwar praktisch nicht bedeutsam, deutet jedoch in eine andere Richtung als die Kohortentrends in Mathematik und den Naturwissenschaften. Inwiefern sich für verschiedene sprachliche Fächer ähnliche Motivationstrends für Mädchen und Jungen zeigen, sollte mithilfe weiterer Trenddaten untersucht werden. Um auszuschließen, dass es sich bei den Fachunterschieden in den Kohortentrends um ein methodisches Artefakt aufgrund unterschiedlicher Messinstrumente handelt, sollten dabei möglichst identische Itemformulierungen über die Fächer hinweg verwendet werden.

Außerdem wurde in den hier genutzten Daten der IQB-Bildungstrendstudien das Geschlecht der Schülerinnen und Schüler binär erfasst. Vor dem Hintergrund der zunehmenden Bewusstheit um die Vielfalt von Geschlecht ist zukünftig eine differenziertere Erfassung des Geschlechts wünschenswert. Aufgrund der anzunehmenden verhältnismäßig kleinen Population Lernender, die sich nicht mit den binären Geschlechtskategorien „weiblich“ bzw. „männlich“ identifizieren oder sich einem anderen oder keinem Geschlecht zugehörig fühlen, sind Large-scale Assessments wie die IQB-Bildungstrends mit ihren großen Stichprobenumfängen für die Untersuchung dieser Gruppen besonders gut geeignet.

Zudem basieren die Ergebnisse unserer Studie auf Trenddaten mit zwei Erhebungszeitpunkten. Das Fortschreiben der Trends in Leistung und Motivation im Rahmen des Bildungsmonitorings wird zeigen können, ob sich die ungünstigen motivationalen Entwicklungen in Mathematik und in den naturwissenschaftlichen Fächern weiter-

führen. Aufschlussreich wäre in diesem Zusammenhang auch eine systematische Übersichtsarbeit zur Entwicklung schulischer Motivation im Trend in Deutschland, die auf Basis möglichst vieler verfügbarer Datensätze einen möglichst langen Zeitraum abdeckt. Psychometrische Ansätze aus dem Bereich der integrativen Datenanalyse, bei der es um die Verknüpfung und gemeinsame Nutzung multipler Datensätze geht (Curran & Hussong, 2009), könnten dabei helfen, das Problem der Verknüpfbarkeit der Messinstrumente (wenn überhaupt, überschneiden sich nur wenige Items zwischen den Studien) zu lösen. Neben Selbstkonzept und Interesse könnten zudem auch weitere Trenddaten motivationaler und psychologischer Konstrukte in diese Analysen einbezogen werden.

In unserer Studie nutzten wir zudem manifest gebildete Skalen. Auch wenn diese durchaus hohe Reliabilitäten aufweisen, sind sie weiterhin messfehlerbehaftet, was zu einer Verschätzung von Effektstärken (z. B. der Kohortentrends) sowie zu Schwierigkeiten beim Vergleich von Effektstärken führen kann, wenn Skalen über verschiedene Fächer oder über die Kohorten unterschiedlich reliabel sind. In der vorliegenden Studie wurde die Priorität auf die Nutzung von Auswertungsverfahren gesetzt, die die komplexe Datenstruktur der Stichprobe berücksichtigen. Daher wurde das Paket *eatRep* verwendet, das multipel imputierte Datensätze mit Stichprobengewichten und dem *Jackknife*-Verfahren zur Standardfehleranpassung verarbeiten kann. In diesem Zusammenhang musste allerdings auf eine latente Modellierung der motivationalen Variablen verzichtet werden. Um den Messfehler besser zu berücksichtigen, sollten zukünftige Studien zu Kohortentrends nach Möglichkeit latente Variablenmodelle nutzen.

Von den geschlechtsspezifischen Kohortentrends mit Rückgängen der Motivation bei Jungen im MINT-Bereich lässt sich die Empfehlung ableiten, bei der Förderung von mathematisch-naturwissenschaftlichen Selbstkonzepten und Interessen zukünftig nicht nur auf Mädchen zu fokussieren, sondern Jungen ebenfalls einzubeziehen. Auch Lehrkräften und Eltern kommt bei der Reduktion von Geschlechtsunterschieden in schulischer Motivation eine wichtige Rolle zu, da sich ihre geschlechtsspezifischen Überzeugungen oder Stereotype bedeutsam auf die Selbstkonzepte und Interessen ihrer Schülerinnen und Schüler bzw. Kinder auswirken (Kollmayer, Schober & Spiel, 2018; Retelsdorf et al., 2015). Es ist daher sinnvoll, sich der eigenen Geschlechterstereotype bewusst zu werden und ihrem Einfluss auf das eigene Verhalten aktiv entgegenzuwirken. So sollten Lehrkräfte etwa bewusst betonen, dass sie von Mädchen und Jungen in allen Domänen gleich gute Leistungen erwarten (Hartley & Sutton, 2013) und im Unterricht möglichst nur Materialien ohne offen dargestellte Geschlechterstereotypen sowie

eine geschlechtsneutrale Sprache verwenden (Hannover & Wolter, 2021).

## Fazit

Im vorliegenden Beitrag gingen wir der Frage nach, ob sich bei Sekundarschülerinnen und -schülern in Deutschland Kohortentrends in den fachbezogenen Selbstkonzepten und Interessen zeigen und ob sich etwaige geschlechtsspezifische Trends auch nach Berücksichtigung von Leistungsunterschieden zwischen den Erhebungszeitpunkten finden. Die identifizierten negativen Trends im Selbstkonzept und Interesse bei Jungen scheinen auf den mathematisch-naturwissenschaftlichen Bereich beschränkt zu sein, die positiven Entwicklungen im Fach Deutsch sprechen gegen eine globale Deutung von Jungen als Bildungsverlierer im motivationalen Bereich. Für Mädchen zeigten sich keine oder geringfügig positive Veränderungen zwischen den Kohorten. Da diese geschlechtsspezifischen Trends in allen untersuchten Fächern nicht bedeutsam auf Testleistungs- und Notenunterschieden zwischen den Erhebungszeitpunkten zurückgeführt werden konnten, müssen in zukünftigen Studien weitere potentiell relevante Variablen sowie gegebenenfalls deren Interaktion für die sich verändernde Motivation – insbesondere von Jungen – untersucht werden. Ziel sollte es sein, Wirkmechanismen zu identifizieren und entsprechende Gegenmaßnahmen zu implementieren, um eine hohe schulische Motivation bei Mädchen und Jungen gleichermaßen zu sichern.

## Elektronisches Supplement (ESM)

Das elektronische Supplement ist mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1024/1010-0652/a000346>

**ESM 1.** Die Datei enthält eine Tabelle zu bisherigen Studien zu (geschlechtsspezifischen) Kohortentrends in Motivation und Leistung (Tabelle E1)

**ESM 2.** Die Datei enthält Tabellen zur Stichprobenbeschreibung und den Reliabilitäten, zu den Korrelationen zentraler Variablen und den Mittelwertvergleichen (Tabellen E1–3)

**ESM 3.** Diese Datei enthält Abbildungen der schulfachbezogenen Selbstkonzepte und Interessen für Mädchen und Jungen pro Notenstufe bzw. Kompetenzstufe nach Erhebungszeitpunkt (Abbildungen E1–9)

**ESM 4.** Diese Datei enthält Tabellen mit Ergebnissen der Regressionsanalysen (M0 und M3) im Rahmen der Robustheitschecks (Tabellen E1–4)

## Literatur

- Arens, A. K. & Jansen, M. (2016). Self-concepts in reading, writing, listening, and speaking: A multidimensional and hierarchical structure and its generalizability across native and foreign languages. *Journal of Educational Psychology*, 108 (5), 646–664.
- Artelt, C., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–112). Münster: Waxmann.
- Bandura, A. & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice Hall: Englewood cliffs.
- Becker, B., Weirich, S., Mahler, N. & Sachse, K. A. (2019). Testdesign und Auswertung des IQB-Bildungstrends 2018: Technische Grundlagen. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018* (S. 411–425). Münster: Waxmann.
- Becker, B., Weirich, S. & Sachse, K. (2021). *eatAnalysis: Convenience function for various analysis types* (Version 0.3.3). Verfügbar unter <https://github.com/beckerbenj/eatAnalysis>
- Böhme, K., Sebald, S., Weirich, S. & Stanat, P. (2016). Geschlechtsspezifische Disparitäten. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015* (S. 377–407). Münster: Waxmann.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Curran, P. J. & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, 14 (2), 81–100.
- Cvencek, D., Meltzoff, A. & Greenwald, A. (2011). Math-gender stereotypes in elementary school children. *Child Development*, 82 (3), 766–779.
- 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.
- Diedrich, J., Schiepe-Tiska, A., Ziernwald, L., Tupac-Yupanqui, A., Weis, M., McElvany, N. & Reiss, K. (2019). Lesebezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen. In K. Reiss, M. Weis, E. Klieme & O. Köller (Hrsg.), *PISA 2018* (S. 81–109). Münster: Waxmann.
- Eccles, J. S. & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task value 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 (1), 109–132.
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, 101859.
- 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 (4), 519–533.
- Gentrup, S. & Stanat, P. (2021). *Jungen als neue Bildungsverlierer? – Zusatzauswertungen des IQB-Ländervergleichs 2012 und des IQB-Bildungstrends 2018 (JuBiv). Schlussbericht*. Humboldt-Universität zu Berlin.
- Gignac, G. E. & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78.
- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zu Ungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25, 89–103.
- Hannover, B. & Wolter, I. (2019). Geschlechterstereotype: Wie sie entstehen und sich auswirken. In B. Kortendiek, B. Riegraf & K. Sabisch (Hrsg.), *Handbuch Interdisziplinäre Geschlechterforschung* (S. 201–210). Wiesbaden: Springer.
- Hannover, B. & Wolter, I. (2021). Schule und Geschlecht. In T. Hascher, T.-S. Idel & W. Helsper (Hrsg.), *Handbuch Schulforschung* (S. 1–21). Wiesbaden: Springer.
- Harel, O. (2009). The estimation of  $R^2$  and adjusted  $R^2$  in incomplete data sets using multiple imputation. *Journal of Applied Statistics*, 36 (10), 1109–1118.
- Hartley, B. L. & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development*, 84, 1716–1733.
- Helbig, M. (2012). Warum bekommen Jungen schlechtere Schulnoten als Mädchen? Ein sozialpsychologischer Erklärungsansatz. *Zeitschrift für Bildungsforschung*, 2, 41–54.
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S. & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73, 509–527.
- Jansen, M., Lüdtke, O. & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, 46, 116–127.
- Jansen, M., Schneider, R., Schipolowski, S. & Henschel, S. (2019). Motivationale Schülermerkmale im Fach Mathematik und in den naturwissenschaftlichen Fächern. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018* (S. 337–354). Münster: Waxmann.
- Jansen, M., Schroeders, U. & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, 30, 11–21.
- Jansen, M., Stäbler, F., Becker, M. & Neumann, M. (2017). Motivationale Orientierungen und Aspekte schulischen Wohlbefindens vor und nach der Berliner Schulstrukturreform. In M. Neumann, M. Becker, J. Baumert, K. Maaz & O. Köller (Hrsg.), *Zweigliedrigkeit im deutschen Schulsystem. Potenziale und Herausforderungen in Berlin* (S. 295–320). Münster: Waxmann.
- 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. & Heyder, A. (2018). Geschlechtsunterschiede. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. Auflage, S. 209–217). Weinheim: Beltz.
- Kessels, U., Rau, M. & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 74 (4), 761–780.
- KMK. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Verfügbar unter [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2015/2015\\_06\\_11-Gesamtstrategie-Bildungsmonitoring.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf)
- Köller, O., Knigge, M. & Tesch, B. (2011). *IQB-Ländervergleich Sprachen 2008/2009 (IQB-LV 2008–9)* (Version 2) [Datensatz]. Berlin: Institut zur Qualitätsentwicklung im Bildungswesen.
- Kollmayer, M., Schober, B. & Spiel, C. (2018). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology*, 15 (4), 361–377.
- Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, 15, 381–395.

- Lorenz, G., Gentrup, S., Kristen, C., Stanat, P. & Kogan, I. (2016). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 68 (1), 89–111.
- Lotz, C., Schneider, R. & Sparfeldt, J. R. (2018). Differential relevance of intelligence and motivation for grades and competence tests in mathematics. *Learning and Individual Differences*, 65, 30–40.
- Mahler, N., Schipolowski, S. & Weirich, S. (2019). Anlage und Durchführung. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018* (S. 99–124). Münster: Waxmann.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76 (2), 397–416.
- Möller J. & Trautwein U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 179–203). Berlin: Springer.
- Muntoni, F. & Retelsdorf, J. (2018). Gender-specific teacher expectations in reading—The role of teachers' gender stereotypes. *Contemporary Educational Psychology*, 54, 212–220.
- 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 (4), 323–345.
- Pant, H. A., Stanat, P., Hecht, M., Heitmann, P., Jansen, M., Lenski, A. E., Penk, C., Pöhlmann, C., Roppelt, A., Schroeders, U. & Siegle, T. (2015). *IQB-Ländervergleich Mathematik und Naturwissenschaften 2012 (IQB-LV 2012)* (Version 4) [Datensatz]. Berlin: Institut zur Qualitätsentwicklung im Bildungswesen.
- Parker, P. D., van Zanden, B. & Parker, R. B. (2018). Girls get smart, boys get smug: Historical changes in gender differences in math, literacy, and academic social comparison and achievement. *Learning and Instruction*, 54, 125–137.
- Retelsdorf, J., Schwartz, K. & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, 107 (1), 186–194.
- Riegle-Crumb, C. & Humphries, M. (2012). Exploring bias in math teachers' perceptions of students' ability by gender and race/ethnicity. *Gender and Society*, 26 (2), 290–322.
- Rost, D. H., Sparfeldt, J. R. & Schilling, S. R. (2007). *DISK-Gitter mit SKSLF-8. Differentielles Schulisches Selbstkonzept-Gitter mit Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten*. Göttingen: Hogrefe.
- Rubie-Davies, C. M. & Peterson, E. R. (2016). Relations between teachers' achievement, over- and underestimation, and students' beliefs for Māori and Pākehā students. *Contemporary Educational Psychology*, 47, 72–83.
- Rubin, D. B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*, 57 (1), 3–18.
- Rüdiger, C., Jansen, M. & Rjosk, C. (2021). „Paul ist nicht so gut in Deutsch“. Geschlechtsspezifische Benotung im Fach Deutsch – eine Sekundäranalyse der Daten des IQB-Bildungstrends 2015. *Psychologie in Erziehung und Unterricht*, 68, 1–17.
- Sachse, K. A., Haag, N. & Weirich, S. (2016). Testdesign und Auswertung des IQB-Bildungstrends 2015: Technische Grundlagen. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015* (S. 509–525). Münster: Waxmann.
- Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197–222). New York, NY: Routledge.
- Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale, NJ: Erlbaum.
- Schiepe-Tiska, A. & Schmidtner, S. (2013). Mathematikbezogene emotionale und motivationale Orientierungen, Einstellungen und Verhaltensweisen von Jugendlichen in PISA 2012. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 99–122). Münster: Waxmann.
- Schiepe-Tiska, A., Simm, I. & Schmidtner, S. (2016). Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 99–132). Münster: Waxmann.
- Schilling, S. R., Sparfeldt, J. R. & Rost, D. H. (2006). Facetten schulischen Selbstkonzepts: Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie*, 20, 9–18.
- Schipolowski, S., Wittig, J., Mahler, N. & Stanat, P. (2019). Geschlechtsspezifische Disparitäten. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018* (S. 237–263). Münster: Waxmann.
- Schmenk, B. (2004). Language learning: A feminine domain? The role of stereotyping in constructing gendered learner identities. *TESOL Quarterly*, 38, 514–524.
- Schurtz, I. M., Pfost, M., Nagengast, B. & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, 34, 32–41.
- Skaalvik, S. & Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles*, 50 (3), 241–252.
- Stanat, P., Böhme, K., Schipolowski, S., Haag, N., Weirich, S., Sachse, K. A., Hoffmann, L. & Federlein, F. (2018). *IQB-Bildungstrend Sprachen 2015 (IQB-BT 2015)* (Version 5) [Datensatz]. Berlin: Institut zur Qualitätsentwicklung im Bildungswesen.
- Stanat, P., Schipolowski, S., Mahler, N., Weirich, S. & Henschel, S. (Hrsg.). (2019). *IQB-Bildungstrend 2018*. Münster: Waxmann.
- 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.
- Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39, 111–133.
- Van Buuren, S. & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45 (3), 1–67.
- Wang, M.-T. & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, 33 (4), 304–340.
- Weirich, S., Hecht, M. & Becker, B. (2021). *eatRep: Educational Assessment Tools for Replication Methods* (Version 0.13.5).
- Weirich, S., Hecht, M., Becker, B. & Zitzmann, S. (2020). Comparing group means with the total mean in random samples, surveys and large-scale assessments: A tutorial and software illustration [Preprint]. Verfügbar unter <https://tinyurl.com/yfc3nhpp>
- 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 (2), 101–120.
- Zoch, G. & Schober, P. S. (2018). Public child-care expansion and changing gender ideologies of parents in Germany. *Journal of Marriage and Family*, 80, 1020–1039.

## Historie

Manuskript eingereicht: 17.09.2021

Manuskript nach Revision angenommen: 25.03.2022

Onlineveröffentlichung: 19.04.2022

**Danksagung**

Wir danken Dr. Sebastian Weirich und Benjamin Becker für die Unterstützung bei der Datenaufbereitung und dem Forschungsdatenzentrum am IQB für die Bereitstellung der Datensätze.

**Autorenschaft**

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

**Förderung**

Das diesem Beitrag zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JQ2001 gefördert.

Open-Access-Veröffentlichung ermöglicht durch die Humboldt-Universität zu Berlin.

**ORCID**

 Rebecca Schneider

<https://orcid.org/0000-0002-8496-4178>

 Sarah Gentrup

<https://orcid.org/0000-0002-0547-3363>

 Malte Jansen

<https://orcid.org/0000-0001-7081-6505>

 Petra Stanat

<https://orcid.org/0000-0003-2843-9050>

**Dr. Sarah Gentrup**

Humboldt-Universität zu Berlin

Unter den Linden 6

10099 Berlin

Deutschland

[sarah.gentrup@hu-berlin.de](mailto:sarah.gentrup@hu-berlin.de)