

Einschätzungen der schulischen Motivation und des Arbeitsverhaltens durch Grundschullehrkräfte und deren Bedeutung für Verzerrungen in Leistungserwartungen

Sarah Gentrup  · Camilla Rjosk · Petra Stanat · Georg Lorenz

Online publiziert: 26. Januar 2018

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2018

Zusammenfassung Verschiedene Studien haben gezeigt, dass Leistungserwartungen von Lehrkräften teilweise in Abhängigkeit vom ethnischen Hintergrund, sozialen Hintergrund sowie vom Geschlecht der Schülerinnen und Schüler stereotyp verzerrt sein können. Im vorliegenden Beitrag wird die Annahme untersucht, dass diese systematischen Verzerrungen in Leistungserwartungen im sprachlichen und mathematischen Bereich teilweise auf die Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens der Schülerinnen und Schüler zurückgeführt werden können. Die Ergebnisse, die auf einer Stichprobe von $N=901$ Kindern aus $N=66$ ersten Grund-

Das diesem Beitrag zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter den Förderkennzeichen 01JC1117A (Otto-Friedrich-Universität Bamberg), 01JC1117B (Universität Mannheim) und 01JC1117C (Humboldt-Universität zu Berlin) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s11618-018-0806-2>) enthalten.

S. Gentrup (✉)

Institut für Erziehungswissenschaften, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Deutschland
E-Mail: sarah.gentrup@hu-berlin.de

Dr. C. Rjosk · Prof. Dr. P. Stanat · Dr. G. Lorenz

Institut zur Qualitätsentwicklung im Bildungswesen (IQB) - Wissenschaftliche Einrichtung der Länder an der Humboldt-Universität zu Berlin e.V., Unter den Linden 6, 10099 Berlin, Deutschland

Dr. C. Rjosk

E-Mail: camilla.rjosk@iqb.hu-berlin.de

Prof. Dr. P. Stanat

E-Mail: iqboffice@iqb.hu-berlin.de

Dr. G. Lorenz

E-Mail: georg.lorenz@iqb.hu-berlin.de

schulklassen mit 69 Klassen- und Fachlehrkräften basieren, unterstützen dies für soziale und geschlechtsbezogene Verzerrungen. Verzerrungen in Abhängigkeit vom ethnischen Hintergrund der Kinder scheinen hingegen unabhängig von der Lehrkräfteinschätzung der Motivation und des Arbeitsverhaltens der Schülerinnen und Schüler aufzutreten. Rückschlüsse auf stereotype Annahmen und deren Bedeutung für unterrichtliches Handeln sowie für Bildungsgleichheit werden diskutiert.

Schlüsselwörter Leistungsmotivation · Arbeitsverhalten · Lehrererwartungen · Lehrerurteile · Stereotype

Teachers' perceptions of students' motivation and learning behaviour and their role in biased teacher achievement expectations

Abstract Several studies have shown that teachers' achievement expectations for their students may be biased to some extent by students' ethnic or social background as well as gender. The present study investigates whether these biases in teachers' expectations of students' performance in early language and mathematics are partly explained by differences in teachers' perceptions of students' motivation and learning behaviour. The analyses are based on data from a sample of $N=901$ first-grade students from $N=66$ classrooms with 69 teachers. The results support the hypothesis for biases associated with social background and gender. Ethnic bias, however, occurred independently of differences in perceived motivation and learning behaviour. We discuss the relevance of stereotypical assumptions for teaching as well as educational disparities.

Keywords Achievement motivation · Learning behaviour · Stereotypes · Teacher expectations · Teacher judgements

1 Einleitung

Die Existenz ethnischer, sozialer und geschlechtsbezogener Bildungsungleichheiten ist empirisch gut belegt (vgl. z.B. Prenzel et al. 2013). Zur Erklärung dieser Disparitäten werden in der empirischen Bildungsforschung verschiedene Faktoren diskutiert, wie etwa Struktur- und Prozessmerkmale der Familien (Baumert et al. 2003), Sprachkenntnisse der Schülerinnen und Schüler (Heinze et al. 2007) sowie verschiedene Arten von Diskriminierung (Diehl und Fick 2016). Eine weitere Annahme bezieht sich auf Leistungserwartungen von Lehrkräften, die für Bildungsgleichheit bedeutsam sein können, da sie eine wichtige Grundlage für pädagogisches Handeln im Unterricht bilden (Schrader und Helmke 2001).

Leistungserwartungen von Lehrkräften basieren nur zum Teil auf tatsächlichen Leistungen der Lernenden (z. B. Südkamp et al. 2012) und scheinen auch unter Kontrolle der Leistungen in Abhängigkeit vom ethnischen Hintergrund, sozialen Hintergrund oder Geschlecht der Heranwachsenden unterschiedlich auszufallen (z. B. Jussim und Harber 2005; Ready und Wright 2011; Lorenz et al. 2016). Obwohl solche systematischen Abweichungen in Leistungserwartungen von Lehrkräften, die

als Verzerrungen bezeichnet werden können, sowohl im internationalen als auch im deutschen Kontext wiederholt belegt wurden, ist über ihre Ursachen wenig bekannt. Angenommen werden kann, dass Lehrkräfte in der Wahrnehmung einzelner Kinder unbewusst und automatisch auf Stereotype, also auf verallgemeinernde Vorstellungen über Merkmale, Eigenschaften und Verhaltensweisen von Schülergruppen zurückgreifen (vgl. Fiske und Neuberg 1990; Lorenz 2017). Stereotype von Lehrkräften können sich dabei auf verschiedene Inhalte beziehen. Denkbar wären beispielsweise verallgemeinernde Vorstellungen über verschiedene Fähigkeiten von Lernenden oder über die häusliche Unterstützung, die bestimmte Gruppen von Schülerinnen und Schülern erfahren. Studien zu Übergangsempfehlungen zeigen allerdings, dass Lehrkräfte neben schulischen Leistungen primär der Motivation und dem Arbeitsverhalten von Lernenden hohe Relevanz für zukünftige Leistungen beimessen (z. B. Böhmer et al. 2015). Als mögliche Quelle für Verzerrungen von Leistungserwartungen kommen daher insbesondere stereotypenbasierte Annahmen über die Motivation und das Arbeitsverhalten der Kinder in Betracht. Der vorliegende Beitrag untersucht deshalb, ob ethnische, soziale und geschlechtsbezogene Verzerrungen in Leistungserwartungen auf Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens ihrer Schülerinnen und Schüler zurückgeführt werden können. Hinausgehend über viele der bisherigen Studien werden dabei in der vorliegenden Arbeit Selbsteinschätzungen der Kinder zu ihrer Motivation und ihrem Arbeitsverhalten berücksichtigt und als Vergleichsmaß für die Lehrkräfteeinschätzungen herangezogen. Anhand dieses Vorgehens lässt sich bestimmen, ob die Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens primär die Selbsteinschätzungen der Lernenden oder eher stereotypenbasierte Annahmen widerspiegeln.

2 Theorie und Forschungsstand

2.1 Leistungserwartungen von Lehrkräften und deren Verzerrungen

Bei Leistungserwartungen von Lehrkräften handelt es sich um Prognosen der zukünftigen Leistungen von Schülerinnen und Schülern. Auch wenn sie sich konzeptionell von Leistungseinschätzungen aktueller Fähigkeiten abgrenzen lassen, werden in der Forschung häufig Einschätzungen aktueller Leistungen als Indikatoren für Leistungserwartungen herangezogen und beide Begriffe synonym gebraucht. Um die Lesbarkeit zu erhöhen, wird im Folgenden zusammenfassend von *Leistungserwartungen* gesprochen. Ein weiteres verwandtes Konzept ist die diagnostische Kompetenz von Lehrkräften. Hiermit ist die übergeordnete Fähigkeit gemeint, zu genauen Urteilen über Merkmale von Lernenden, wie z. B. Leistungen, zu gelangen (vgl. Schrader 2010). Um den Forschungsstand umfassend darzustellen, werden im Folgenden einschlägige Ergebnisse aller drei Konzeptionen berücksichtigt. Für die Operationalisierung der Leistungserwartungen werden in der vorliegenden Arbeit jedoch ausschließlich Prognosen zukünftiger Leistungen der Schülerinnen und Schüler herangezogen.

Zahlreiche Studien zeigen, dass Leistungserwartungen von Lehrkräften zu einem hohen Anteil die tatsächlichen Leistungen von Heranwachsenden widerspiegeln

(Jussim et al. 1996). Laut Metaanalysen beträgt die geteilte Varianz von Leistungserwartungen und standardisiert erfassten Leistungen etwa 40% (Hoge und Coladarci 1989; Südkamp et al. 2012). Dies bedeutet jedoch auch, dass ein substanzieller Varianzanteil der Erwartungen nicht auf die Leistungen zurückgeführt werden kann.

Verschiedene Studien verdeutlichen zudem, dass die beobachteten Über- und Unterschätzungen nicht rein zufällig sind. Vielmehr traten über Leistungsunterschiede hinaus systematische Unterschiede in den Leistungserwartungen von Lehrkräften für verschiedene Schülergruppen auf (vgl. auch Lorenz et al. 2016). Derartige Unterschiede werden im Folgenden als *Verzerrungen* bezeichnet.

Internationale Studien weisen darauf hin, dass Leistungserwartungen von Lehrkräften an Kinder aus zugewanderten Familien (z. B. Jussim und Harber 2005; McKown und Weinstein 2008) und an Kinder aus sozial benachteiligten Familien negativ verzerrt sind (z. B. Jussim et al. 1996; Ready und Wright 2011). Die Befunde zu geschlechtsbezogenen Verzerrungen sind hingegen uneinheitlich und deuten auf bereichsspezifische Effekte hin. So fanden einige Studien eine Überschätzung von Jungen im mathematischen (Jussim und Eccles 1992) und von Mädchen im sprachlichen Bereich (Hinnant et al. 2009; Ready und Wright 2011). Andere Untersuchungen identifizierten keine Verzerrungen in Abhängigkeit vom Geschlecht (McKown und Weinstein 2008; van den Bergh et al. 2010).

In Deutschland wurden bisher kaum Studien zu Verzerrungen in Leistungserwartungen durchgeführt. In Untersuchungen zur diagnostischen Kompetenz von Lehrkräften zeigten sich nach Kontrolle standardisierter Leistungsmaße keine Unterschiede in den Leistungsurteilen für Kinder mit und ohne Zuwanderungshintergrund (Schrader und Helmke 1990; Stahl 2007; Hachfeld et al. 2010; Rjosk et al. 2011). Diese Analysen differenzierten jedoch nicht nach ethnischen Herkunftsgruppen, für die unterschiedliche Stereotype und darauf basierende Annahmen bestehen könnten.

Zu sozialen und geschlechtsbezogenen Verzerrungen in Leistungsurteilen sind die Befunde im deutschsprachigen Raum uneinheitlich. In einigen Studien wurden keine sozialen (Hachfeld et al. 2010; Karing et al. 2011) oder geschlechtsbezogenen Verzerrungen (Schrader und Helmke 1990; Karing et al. 2011) identifiziert. Stahl (2007) berichtete hingegen Überschätzungen im Bereich Lesen für Kinder aus sozial bessergestellten Familien und für Mädchen.

Aufgrund der uneinheitlichen und insgesamt unzureichenden Befundlage, sind Lorenz et al. (2016) anhand von Daten der in Deutschland durchgeführten KuL-Studie der Frage nachgegangen, inwieweit Leistungserwartungen von Lehrkräften in Abhängigkeit vom ethnischen und sozialen Hintergrund sowie vom Geschlecht der Lernenden verzerrt sind. Im Fach Deutsch erwarteten die Lehrkräfte bei gleichen Testleistungen der Erstklässlerinnen und Erstklässler eine weniger günstige Leistungsentwicklung für türkischstämmige Kinder im Vergleich zu Kindern ohne Zuwanderungshintergrund, für Kinder aus sozial schwächeren Familien im Vergleich zu Kindern aus sozial bessergestellten Familien sowie für Jungen im Vergleich zu Mädchen. Im Fach Mathematik fielen die Leistungserwartungen bei gleichen tatsächlichen Leistungen höher aus für Kinder mit osteuropäischem Hintergrund als für Kinder ohne Zuwanderungshintergrund und höher für Kinder aus sozial bessergestellten Familien als für Kinder aus sozial benachteiligten Familien. Geschlechtsbezogene Verzerrungen waren im Fach Mathematik nicht zu verzeichnen.

Zusammenfassend scheinen Leistungserwartungen von Lehrkräften nicht nur die tatsächlichen Leistungen ihrer Schülerinnen und Schüler widerzuspiegeln. Außerdem wird deutlich, dass Leistungserwartungen in Abhängigkeit von Hintergrundmerkmalen der Lernenden verzerrt sein können. Offen ist allerdings, worauf diese Verzerrungen zurückzuführen sind und ob ihnen stereotype Annahmen über weitere lernrelevante Merkmale der Lernenden, wie etwa deren Motivation und Arbeitsverhalten, zugrunde liegen.

2.2 Stereotype Wahrnehmungsprozesse

Als theoretischer Rahmen zur Erklärung systematischer Verzerrungen in Leistungserwartungen eignet sich das sozialpsychologische Kontinuum-Modell (Fiske und Neuberg 1990). Diesem Modell zufolge bewegt sich die Personenwahrnehmung zwischen zwei Polen – der individualisierenden und der kategorisierenden Verarbeitung. Bei letzterer werden Personen anhand eines einfach zu identifizierenden Merkmals einer sozialen Kategorie zugeordnet (Ganter 1997), wobei die Merkmale Alter, Geschlecht, ethnische Herkunft sowie sozialer Status besonders salient sind (Fiske und Neuberg 1990). Es wird angenommen, dass die Zuordnung einer Person zu einer Gruppe kognitive Stereotype aktiviert, also verallgemeinernde Vorstellungen über Merkmale, Eigenschaften und Verhaltensweisen der Gruppenmitglieder (Ganter 1997). Im Wahrnehmungsprozess findet im Allgemeinen zuerst eine kategorisierende Verarbeitung statt. Eine Verschiebung hin zur individualisierenden Verarbeitung, bei der individuelle Eigenschaften und Merkmale der jeweiligen Person berücksichtigt werden, erfolgt erst unter bestimmten Voraussetzungen wie z. B. dem Vorliegen stereotypen-inkonsistenter Informationen (Fiske und Neuberg 1990; Ganter 1997). Ohne eine derartige Verschiebung wird die weitere Wahrnehmung durch Stereotype gelenkt.

Auch in schulischen Situationen können stereotype Wahrnehmungsprozesse auftreten, etwa wenn Einschätzungen von Lehrkräften durch stereotype Merkmalsannahmen für bestimmte Schülergruppen gefärbt sind (Hirschauer und Kullmann 2010). Diese Annahmen können sich auf unterschiedliche Inhalte beziehen. Anhaltspunkte dafür, welche Merkmale Lehrkräfte für Beurteilungen von Schülerinnen und Schülern als bedeutsam erachten, liefern zum Beispiel Interview- und MouseLab-Studien, die Übergangsempfehlungen untersuchten. Neben schulischen Leistungen der Heranwachsenden sprechen Lehrkräfte diesen Studien zufolge der Motivation und dem Arbeitsverhalten (z. B. Ausdauer/ Fleiß, Interesse/ Motivation) die höchste Relevanz zu. Zusätzlich griffen die Lehrkräfte auf Informationen zur elterlichen Unterstützung zurück. Weiteren Merkmalen wie beispielsweise dem Sozialverhalten der Kinder oder dem Beruf der Eltern wurde eine vergleichsweise geringere Bedeutung zugesprochen (Nölle et al. 2009; Böhmer et al. 2015). Dass die Motivation und das Arbeitsverhalten von Lernenden tatsächlich prädiktiv für Schulleistungen und weiteres Lernen sind (vgl. z. B. Spinath 2011), spricht ebenfalls für die Annahme, dass Lehrkräfte motivationalen Merkmalen bei der Prognose von Leistungsentwicklungen hohe Relevanz beimessen und sie in ihre Leistungserwartungen einbeziehen. Auch Studienergebnisse zu Zusammenhängen von Leistungserwartungen und Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens legen nahe, dass dieser

Einbezug tatsächlich stattfindet: Ein Kind, dessen Motivation und Arbeitsverhalten von einer Lehrkraft als vergleichsweise günstig eingeschätzt wird, scheint bei gleichen Leistungen auch als begabter und leistungsstärker wahrgenommen zu werden (Schrader und Helmke 1990; Anders et al. 2010), bessere Noten zu erhalten (Trautwein und Baeriswyl 2007; Anders et al. 2010; Kuhl und Hannover 2012) und eher für den Übergang auf ein Gymnasium empfohlen zu werden (Anders et al. 2010).

Legt man auch für die Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens von Schülerinnen und Schülern das Kontinuum-Modell der Personenwahrnehmung zugrunde, existieren zwei Wege, über die die Lehrkräfte zu ihren Einschätzungen gelangen können: Zum einen über die Verarbeitung individueller Informationen, die ihnen über die Kinder vorliegen, und zum anderen über die Anwendung stereotyper Annahmen über die Motivation und das Arbeitsverhalten für bestimmte Gruppen. Eine Verarbeitung anhand von Stereotypen ist durchaus wahrscheinlich, da einige Studien zeigen, dass Lehrkräfte die Lernmotivation (Spinath 2005) sowie verwandte emotional-motivationale Merkmale wie z. B. das schulische Interesse (Karing 2009) oder lernbezogene Emotionen (Givvin et al. 2001) nur relativ ungenau einschätzen. Einschätzungen der Motivation und des Arbeitsverhaltens könnten daher ebenfalls anfällig für Einflüsse stereotyper Personenwahrnehmung sein.

Inwieweit Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens systematisch mit Hintergrundmerkmalen der Schülerinnen und Schüler variieren und ob diese möglichen Variationen Unterschiede in den Selbsteinschätzungen der Lernenden widerspiegeln, soll im Folgenden dargestellt werden.

2.3 Unterschiede in Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens in Abhängigkeit vom ethnischen Hintergrund, sozialen Hintergrund und Geschlecht der Schülerinnen und Schüler

Bislang haben nur wenige Studien untersucht, ob Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens von Lernenden systematisch mit Hintergrundmerkmalen wie der ethnischen und sozialen Herkunft oder dem Geschlecht der Schülerinnen und Schüler in Verbindung stehen und die entsprechenden Befunde sind uneinheitlich. Außerdem wurden Selbsteinschätzungen der Schülerinnen und Schüler in den meisten Studien nicht berücksichtigt. Nur drei Studien liefern bislang Anhaltspunkte dafür, inwieweit die Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens den Selbsteinschätzungen der Schülerinnen und Schüler entsprechen oder ob sie eher auf stereotypen Annahmen über diese Merkmale basieren.

Bezogen auf den *ethnischen Hintergrund* Heranwachsender zeigte eine experimentelle Studie aus den USA, dass Grundschullehrkräfte Jungen mit kaukasisch-amerikanisch klingendem Namen als schulisch motivierter beschrieben als Jungen mit afro-amerikanisch klingendem Namen (Anderson-Clark et al. 2008). In einer ebenfalls in den USA durchgeführten Feldstudie schätzten Lehrkräfte Heranwachsende mit und ohne Zuwanderungshintergrund dagegen ähnlich anstrengungsbereit ein, was mit den Selbsteinschätzungen der Schülerinnen und Schüler übereinstimmte (Jussim et al. 1996).

In Deutschland wurden bisher in zwei Studien Lehrkräfteeinschätzungen zu motivationalen Merkmalen und Aspekten des Arbeitsverhaltens für Heranwachsende mit und ohne Zuwanderungshintergrund analysiert. In diesen Studien nahmen Lehrkräfte für Kinder ohne Zuwanderungshintergrund eine etwas höhere Motivation und ein etwas günstigeres Arbeitsverhalten (Anders et al. 2010) sowie eine leicht höhere Schulfreude, Anstrengungsbereitschaft und günstigere schulische Arbeitshaltung an (Schneider 2011) als für Kinder aus zugewanderten Familien, wobei jedoch keine Selbsteinschätzungen der Kinder kontrolliert wurden. Ergebnisse für verschiedene ethnische Gruppen in Deutschland liegen bisher ebenfalls nicht vor.

Auch in Abhängigkeit von der *sozialen Herkunft* der Heranwachsenden identifizierten verschiedene Studien Unterschiede in Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens. So wurden Kinder aus sozial bessergestellten Familien nicht nur in Deutschland (Trautwein und Baeriswyl 2007; Anders et al. 2010) sondern auch in Studien aus den USA (Jussim et al. 1996) tendenziell als motivierter und anstrengungsbereiter eingeschätzt als Kinder aus sozial benachteiligten Familien. Diese von den Lehrkräften angenommenen Gruppenunterschiede fielen dabei stärker aus als die entsprechenden Gruppenunterschiede in den Selbsteinschätzungen der Kinder (Jussim et al. 1996).

Darüber hinaus weisen mehrere Studienergebnisse darauf hin, dass Lehrkräfte die Motivation und das Arbeitsverhalten in Abhängigkeit vom *Geschlecht* der Kinder unterschiedlich einschätzen. So scheinen Lehrkräfte davon auszugehen, dass Mädchen im Vergleich zu Jungen insgesamt motivierter sind (Trautwein und Baeriswyl 2007; Anders et al. 2010; Schneider 2011), sich stärker anstrengen (Jussim und Eccles 1992; Jussim et al. 1996) sowie über eine höhere schulische Selbstregulationskompetenz (Duckworth und Seligman 2006; Kuhl und Hannover 2012) verfügen. Inwieweit diese angenommenen Geschlechtsunterschiede Differenzen in Selbsteinschätzungen der jeweiligen Schülerinnen und Schüler widerspiegeln, bleibt allerdings uneindeutig. Lediglich drei Studien haben zusätzlich zu den Einschätzungen der Lehrkräfte auch Selbsteinschätzungen der Schülerinnen und Schüler berücksichtigt. Diese Studien kamen zu uneinheitlichen Ergebnissen: In der Studie von Duckworth und Seligman (2006) stimmten die von den Lehrkräften angenommenen Geschlechtsunterschiede in der Selbstregulationskompetenz mit den Selbsteinschätzungen der Mädchen und Jungen überein. In den Studien von Jussim und Eccles (1992) und Jussim et al. (1996) wurden hingegen Differenzen zwischen den Einschätzungen der Lehrkräfte und den Angaben der Mädchen und Jungen dazu deutlich, wie sehr sie sich in der Schule anstrengen und wie viel Zeit sie in Hausaufgaben investieren.

In der Zusammenschau deutet die Befundlage darauf hin, dass Lehrkräfte die Motivation und das Arbeitsverhalten von Schülerinnen und Schülern in Abhängigkeit von deren Hintergrundmerkmalen unterschiedlich einschätzen. Da sich diese Unterschiede ersten Vergleichen zufolge nur teilweise in den Selbsteinschätzungen der Kinder widerspiegeln scheinen, könnte die Wahrnehmung dieser Unterschiede teilweise auf stereotypen Annahmen der Lehrkräfte beruhen. Aufgrund der unzureichenden und uneinheitlichen Befundlage bedarf diese Hypothese jedoch weiterer empirischer Analysen.

Sollten Lehrkräfte Unterschiede in der Motivation und dem Arbeitsverhalten für verschiedene Gruppen von Schülerinnen und Schülern in Abhängigkeit von deren

ethnischem Hintergrund, sozialem Hintergrund oder Geschlecht annehmen, die sich nicht in deren tatsächlichen Motivlagen und Verhalten widerspiegeln, kann es dadurch zu den beobachteten Verzerrungen in Leistungserwartungen kommen.

2.4 Vermittlung von Verzerrungen in Leistungserwartungen durch die Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens

Die Frage, ob die von Lehrkräften eingeschätzte Motivation und das eingeschätzte Arbeitsverhalten der Schülerinnen und Schüler ethnische, soziale und geschlechtsbezogene Verzerrungen in ihren Leistungserwartungen erklären können, wurde bislang noch nicht empirisch untersucht. Zwei Studien haben allerdings die Mediation des Zusammenhangs zwischen Geschlecht und Noten (als Indikatoren für Leistungseinschätzungen) durch die von der Lehrkraft eingeschätzte Selbstregulationskompetenz der Kinder analysiert und bieten damit erste Hinweise. Die Studien zeigen, dass die von den Lehrkräften als höher eingeschätzte Selbstregulationskompetenz von Mädchen den Zusammenhang zwischen Geschlecht und Noten vermittelte (Duckworth und Seligman 2006; Kuhl und Hannover 2012). In getrennten Analysen für die Deutsch- und Mathematiknoten wurden zudem unterschiedliche Mediationsmuster in den Fächern deutlich (Kuhl und Hannover 2012): Mädchen erhielten im Fach Deutsch auch bei gleicher Leseleistung bessere Noten als Jungen, wobei dieser Zusammenhang größtenteils durch die von den Lehrkräften eingeschätzte Selbstregulationskompetenz vermittelt wurde. Im Gegensatz hierzu waren die Noten der Mädchen und Jungen im Fach Mathematik allgemein und auch unter Kontrolle der Testleistung vergleichbar. Wurde jedoch zusätzlich die Lehrkrifteinschätzung der Selbstregulationskompetenz berücksichtigt, erhielten Jungen im Durchschnitt bessere Mathematiknoten als Mädchen.

Diese Befunde deuten darauf hin, dass Geschlechtsunterschiede in der Benotung durch Lehrkräfte auf deren Annahmen zur Selbstregulationskompetenz, die eine wichtige Voraussetzung für effektives Arbeitsverhalten darstellt, zurückgeführt werden können. Unklar ist allerdings, ob auch ethnische, soziale und geschlechtsbezogene Verzerrungen in Leistungserwartungen über die Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens vermittelt werden.

3 Forschungsfragen und Hypothesen

Der vorliegende Beitrag knüpft an den Analysen von Lorenz et al. (2016) an, in denen anhand von Daten des KuL-Projekts gezeigt werden konnte, dass die Leistungserwartungen von Lehrkräften in Abhängigkeit vom ethnischen Hintergrund, sozialen Hintergrund und Geschlecht der Kinder verzerrt sind. Anhand desselben Datensatzes gehen die im Folgenden beschriebenen Analysen der Frage nach, ob sich diese Verzerrungen der Leistungserwartungen auf Unterschiede in der Lehrkrifteinschätzung der schulischen Motivation und des Arbeitsverhaltens ihrer Schülerinnen und Schüler zurückführen lassen.

Dazu wird zunächst untersucht, ob sich die Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens in Abhängigkeit von den Hintergrundmerkmalen der

Kinder allgemein und nach Kontrolle der Selbsteinschätzungen der Lernenden unterscheiden. Da der theoretische und empirische Forschungsstand zu ethnischen Unterschieden nicht ausreicht, um konkrete Annahmen für einzelne Herkunftsgruppen abzuleiten, wird die entsprechende Hypothese ungerichtet formuliert.

Hypothese 1a Lehrkräfte schätzen die Motivation und das Arbeitsverhalten von Kindern in Abhängigkeit vom ethnischen Hintergrund unterschiedlich ein.

Hypothese 1b Lehrkräfte schätzen die Motivation und das Arbeitsverhalten von Kindern aus sozial bessergestellten Familien positiver ein als von Kindern aus sozial schwächeren Familien.

Hypothese 1c Lehrkräfte schätzen die Motivation und das Arbeitsverhalten von Mädchen positiver ein als von Jungen.

Es wird erwartet, dass die in den Hypothesen 1a–c postulierten Unterschiede auch nach Kontrolle der Selbsteinschätzungen der Motivation und des Arbeitsverhaltens der Schülerinnen und Schüler bestehen bleiben.

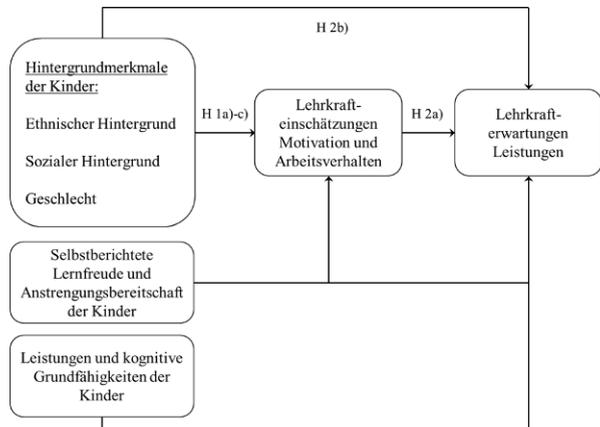
Weiterhin wird der Frage nachgegangen, ob Unterschiede in der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens Verzerrungen in den Leistungserwartungen erklären.

Hypothese 2a Zwischen Leistungserwartungen von Lehrkräften und ihrer Einschätzung der Motivation und des Arbeitsverhaltens der Kinder besteht ein positiver Zusammenhang.

Hypothese 2b Die Zusammenhänge zwischen den Hintergrundmerkmalen der Lernenden und den Leistungserwartungen der Lehrkräfte lassen sich zum Teil auf die Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens zurückführen, d. h. die Koeffizienten des ethnischen Hintergrunds, sozialen Hintergrunds und Geschlechts in Prädiktion der Leistungserwartungen reduzieren sich nach Kontrolle der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens ihrer Schülerinnen und Schüler signifikant.

Eine grafische Darstellung des vollständigen Analysemodells ist in Abb. 1 zu finden.

Abb. 1 Grafische Darstellung des vollständigen Analysemodells und der Hypothesen (H 1a bis H 2b)



4 Methodisches Vorgehen

4.1 Stichprobe

Die vorliegende Untersuchung basiert auf Daten der Studie „Kompetenzerwerb und Lernvoraussetzungen (KuL)“¹, an der im Schuljahr 2013/2014 erste Grundschulklassen aus $N=38$ Schulen aus Nordrhein-Westfalen (NRW) auf freiwilliger Basis teilnahmen. Die Rücklaufquoten der Einwilligungen zur Studienteilnahme lagen auf Schulebene bei 8 %, auf Lehrkräfebene bei 74 % und auf Eltern/ Kind-Ebene bei 66 %. In die folgenden Analysen gehen Daten von insgesamt $N=901$ Kindern aus $N=66$ ersten Klassen mit 69 Klassen- und Fachlehrkräften ein.

Die Lehrkräfte waren im Durchschnitt 41 Jahre und 8 Monate alt ($SD=8,89$) und seit durchschnittlich 12 Jahren ($SD=9,00$) im Lehrberuf tätig. Es handelte sich fast ausschließlich um weibliche Lehrkräfte (94 %) ohne Zuwanderungshintergrund (94 %).

Zum Zeitpunkt der Einschulung waren die Kinder (48 % weiblich) im Mittel 6 Jahre und 5 Monate alt ($SD=0,32$). Laut Elternaussagen zum Geburtsland beider Elternteile sowie der Großeltern lag der Anteil von Kindern mit Zuwanderungshintergrund bei 38 % (mindestens ein Großeltern- oder Elternteil nicht in Deutschland geboren). Kinder der ersten Zuwanderungsgeneration ($N=21$)² wurden ebenso wie Kinder ohne Angaben zum Zuwanderungshintergrund ($N=134$) aus den Analysen ausgeschlossen. Die Angaben zum Herkunftsland der Familien wurden zu insgesamt vier Herkunftsregionen zusammengefasst: „Deutschland“ (61,6 %), „Türkei“ (11,3 %), „ehem. SU und Osteuropa“ (10,9 %) und „andere Länder“ (16,2 %), wobei

¹ Das interdisziplinäre Forschungsprojekt wurde unter der Leitung von Prof. Dr. Cornelia Kristen (Otto-Friedrich-Universität Bamberg), Prof. Dr. Irena Kogan (Universität Mannheim) und Prof. Dr. Petra Stanat (Humboldt-Universität zu Berlin) durchgeführt.

² Die selbst zugewanderten Erstklässlerinnen und Erstklässler leben erst relativ kurz in Deutschland. Es ist anzunehmen, dass dies Urteile von Lehrkräften in besonderer Weise beeinflusst. Separate Analysen wären wünschenswert, waren durch die geringe Fallzahl jedoch nicht möglich.

die Region „ehem. SU und Osteuropa“³ alle Staaten der ehemaligen Sowjetunion sowie Polen, Tschechien, Bulgarien, Rumänien, Ungarn, Slowakei und Slowenien umfasst.

Als Indikator des sozioökonomischen Status (SES) dient der *Highest International Socio-Economic Index of Occupational Status* (HISEI; Ganzeboom 2010). Die Familien wiesen durchschnittlich einen HISEI von 53,25 ($SD=19,48$) auf. Verglichen mit dem durchschnittlichen HISEI von 48,10 ($SD=14,80$) für Viertklässlerinnen und Viertklässler in NRW, der im IQB-Ländervergleich 2011 ermittelt wurde (Richter et al. 2012), ist die Stichprobe leicht positiv selektiert.

4.2 Datenerhebung und Messinstrumente

In der dritten bis zwölften Unterrichtswoche⁴ nach Schuljahresanfang gaben die Lehrkräfte ihre Einschätzungen der Leistungen, der Motivation und des Arbeitsverhaltens der teilnehmenden Kinder ihrer Klasse in einem Fragebogen an. Im selben Zeitraum wurden die Kompetenzen der Schülerinnen und Schüler in zwei Einzelsitzungen je Kind (jeweils ca. 30 min) von geschulten Testleiterinnen und Testleitern erfasst.

4.2.1 Leistungserwartungen und Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens

Die Lehrkräfte schätzten für jedes Kind die voraussichtliche Leistungsentwicklung im Verlauf des kommenden Schuljahres sowie die zukünftigen Leistungen am Ende des Schuljahres im sprachlichen und mathematischen Bereich auf einer fünfstufigen Antwortskala im Vergleich zu den anderen Kindern der Klasse ein (z. B. „Wo wird das Kind mit seinen Leistungen am Ende des kommenden Schuljahres im Vergleich zu den anderen Kindern aus dieser Klasse voraussichtlich stehen? a) in Mathematik, b) in Deutsch“).⁵ Die Erwartungen für den mathematischen Bereich wurden mit zwei Items ($\alpha=0,94$), die Erwartungen für den sprachlichen Bereich mit drei Items ($\alpha=0,94$) erfasst. Die Items wurden teils aus dem Projekt BiKS-3-10 (Artelt et al. 2013) adaptiert und teils in der KuL-Studie selbst entwickelt.

Zusätzlich wurden die Lehrkräfte gebeten, die Motivation und das Arbeitsverhalten der Kinder auf einer fünfstufigen Skala (1 = *trifft überhaupt nicht zu* bis 5 = *trifft voll und ganz zu*) einzuschätzen. Auch diese Items wurden teils aus dem Projekt BiKS-3-10 (Artelt et al. 2013) adaptiert und teils in der KuL-Studie selbst entwickelt. Von den insgesamt acht Items ($\alpha=0,96$) beziehen sich jeweils drei auf den Deutsch- und Mathematikunterricht (Deutsch: $\alpha=0,92$, Mathematik: $\alpha=0,92$; z. B. „Das Kind hat große Freude am Lernen im Deutsch-/Mathematikunterricht.“) und

³ Zur besseren Lesbarkeit wird nachfolgend für diese Herkunftsregion von „osteuropäischer Hintergrund“ gesprochen. Kinder, deren Familien aus einem nicht-europäischen Land der ehemaligen Sowjetunion zugewandert sind, sind damit gleichermaßen gemeint.

⁴ Der Zeitpunkt, zu dem die Lehrkräfte den Fragebogen zurückschickten, hing nicht mit der Genauigkeit der leistungsbezogenen Erwartungen der Lehrkräfte zusammen (Lorenz et al. 2016).

⁵ Für den genauen Wortlaut aller Items und der Antwortskalen siehe Abb. A1 im Online Anhang.

zwei allgemein auf das Lernen in der Schule (z.B. „Das Kind ist sehr fleißig in der Schule.“). Die Items decken verschiedene motivationale Aspekte und Aspekte des Arbeitsverhaltens ab (vgl. Abb. A1 im Online Anhang). Zur Erhöhung der Vergleichbarkeit zwischen den Lehrkräfteeinschätzungen und den Selbsteinschätzungen der Kinder, die sich fast ausschließlich allgemein auf das Lernen in der Schule beziehen und nicht nach Fächern differenzieren (vgl. Abschn. 4.2.3), wurden die Items in allen dargestellten Analysen als Gesamtmittelwert berücksichtigt. Um die Robustheit der Ergebnisse abzusichern, wurden alle Analysen jedoch auch mit nach Fächern getrennten Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens durchgeführt. Die Ergebnisse fallen sehr ähnlich aus und werden daher nicht dargestellt. Auf bedeutsam abweichende Ergebnisse wird jeweils hingewiesen.

4.2.2 Sprachliche, mathematische und allgemeine kognitive Schülerfähigkeiten

Die mathematischen und sprachlichen (Vorläufer-)Kompetenzen der Kinder wurden mit dem computergestützten Instrument *Fähigkeitsindikatoren Primarschule* (FIPS; Bäuerlein et al. 2012) erhoben. Die Skala *Lautbewusstheit* ($\alpha=0,82$) erfasst Fähigkeiten der phonologischen Bewusstheit. Sie setzt sich aus 26 Aufgaben zusammen, bei denen die Kinder Silben und Fantasiewörter nachsprechen, Reime erkennen und Wörter in ihre einzelnen Laute zerlegen sollen. Die Skala *Lesen* ($\alpha=0,96$) umfasst 88 Aufgaben, die Vorstellungen des Kindes über das Lesen sowie erste Leseaufgaben beinhalten. Der Skala *Mathematik* ($\alpha=0,92$) sind 53 Aufgaben zugeordnet. Nach leichteren Additions- und Subtraktionsaufgaben anhand von Bildern und dem Vorlesen verschiedener Zahlen, erhöht sich der Abstraktionsgrad über Rechenaufgaben anhand von Punkten bis hin zu formalen Rechenaufgaben. Die im Schwierigkeitsgrad aufsteigenden Aufgaben wurden den Kindern in allen Kompetenzbereichen solange dargeboten, bis sie eine bestimmte Anzahl an falschen Antworten gegeben haben. Je nach Kompetenzstand erfasst das Instrument damit bei einigen Kindern vor allem Vorläuferkompetenzen, bei anderen Kindern, die in der Entwicklung bereits fortgeschritten sind, auch sprachliche und mathematische Kompetenzen.

Zusätzlich wurden kognitive Grundfähigkeiten der Kinder erfasst: Zum einen mit dem *Matrizentest des Grundintelligenztests Skala 1* (CFT1; Weiß und Osterland 1997; $\alpha=0,78$) als Maß des schlussfolgernden Denkens und zum anderen mit der *Unterskala Kurzzeitgedächtnis* des FIPS (Bäuerlein et al. 2012; $\alpha=0,76$) als Maß des räumlich-visuellen Arbeitsgedächtnisses.

4.2.3 Schülerelbsteinschätzung der Motivation und des Arbeitsverhaltens

In der Mitte des Schuljahres wurden die Kinder zu ihrer Motivation und zu ihrem Arbeitsverhalten befragt. Dazu wurden zwei Unterskalen des FEES1-2 (Rauer und Schuck 2004) in adaptierter Version eingesetzt. Die Kinder machten auf einer dreistufigen Skala (0=*trifft nicht zu* bis 2=*trifft zu*) Angaben zu ihrer Lernfreude ($\alpha=0,78$, 13 Items; z.B. „Ich lerne gern in der Schule.“) und Anstrengungsbereitschaft ($\alpha=0,70$, 13 Items; z.B. „Ich versuche, auch ganz schwierige Aufgaben zu lösen.“). Die Items beziehen sich größtenteils allgemein auf das Lernen in der Schule.

4.3 Analysestrategie

Zunächst wurde mit Regressionsanalysen untersucht, ob sich die Lehrkrifteinschätzungen der Motivation und des Arbeitsverhaltens in Abhängigkeit von den Hintergrundmerkmalen der Kinder (ethnischer Hintergrund, sozialer Hintergrund, Geschlecht) allgemein unterscheiden (Hypothesen 1a–c). Um spezifische Effekte identifizieren zu können, erfolgte die Aufnahme aller drei Merkmale in das Modell simultan (Tab. 3, Modell 1). Um zu überprüfen, inwieweit die gefundenen Unterschiede mit den tatsächlichen Motivlagen und Arbeitshaltungen der Kinder übereinstimmen oder darüber hinausgehen, wurden die Selbsteinschätzungen der Lernfreude und Anstrengungsbereitschaft in einem nächsten Schritt hinzugenommen (Tab. 3, Modell 2).

Mit Pfadmodellen wurde anschließend analysiert, inwieweit sich Verzerrungen in Leistungserwartungen auf Unterschiede in der Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens der Kinder zurückführen lassen (Hypothesen 2a–b). Als Verzerrungen werden hierbei ethnische, soziale und geschlechtsbezogene Unterschiede in den Leistungserwartungen verstanden, die nach Kontrolle fachspezifischer Leistungen der Kinder, ihrer kognitiven Grundfähigkeiten und ihrer selbsteingeschätzten Lernfreude und Anstrengungsbereitschaft bestehen bleiben (Tab. 4 und 5, Modell 0). Um den Anteil zu bestimmen, mit dem sich Verzerrungen in Leistungserwartungen auf die Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens zurückführen lassen, wurden die Regressionskoeffizienten der Hintergrundmerkmale vor und nach zusätzlicher Kontrolle der Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens miteinander verglichen und indirekte Pfade spezifiziert (Tab. 4 und 5, Modell 1).

Die Modellschätzungen erfolgten mit der Software Mplus Version 7.3 (Muthén und Muthén 1998–2012). Die Schätzung der Standardfehler wurde für die Mehrebenenstruktur der Daten korrigiert und die akzeptierte α -Fehlerwahrscheinlichkeit auf 5 % festgelegt. Alle metrischen Prädiktorvariablen gingen z-standardisiert in die Analysen ein. Die Regressionskoeffizienten wurden an der Varianz der abhängigen Variable standardisiert (y-Standardisierung). Zum Umgang mit fehlenden Werten wurde das in Mplus implementierte Full-Information-Maximum-Likelihood-Verfahren (FIML) unter Einbezug vielfältiger Hilfsvariablen angewendet, u. a. zusätzliche Lehrkrifteinschätzungen, Leistungsmaße sowie Merkmale des familiären Lernumfelds. Der Anteil fehlender Werte in den Analysevariablen lag bei maximal 6,9 % (vgl. Tab. A1 im Online Anhang).

5 Ergebnisse

5.1 Deskriptive Ergebnisse

Mittelwerte und Standardabweichungen der Leistungserwartungen und der Lehrkrifteinschätzung der Motivation und des Arbeitsverhaltens sowohl für die Analysetichprobe insgesamt als auch für die verschiedenen Schülergruppen sind in Tab. 1 zu finden. Tab. 2 enthält Angaben zu den Leistungen und zur selbsteingeschätzten

Lernfreude und Anstrengungsbereitschaft der Lernenden. Die zentralen Ergebnisse sollen im Folgenden kurz beschrieben werden.

Die Mittelwerte der Leistungserwartungen lagen sowohl für den sprachlichen Bereich ($M=3,39$; $SD=0,94$) als auch für den mathematischen Bereich ($M=3,43$; $SD=0,85$) insgesamt knapp oberhalb des theoretischen Mittelpunkts der Skala. Die Lehrkräfte tendierten also insgesamt zu leicht positiven Erwartungshaltungen. Sie gingen demnach davon aus, dass die an der Studie teilnehmenden Kinder ihrer Klasse durchschnittlich eine etwas bessere Leistungsentwicklung zeigen werden als die Kinder der Klasse insgesamt. Bei der Einschätzung der Motivation und des Arbeitsverhaltens war die Tendenz zu Urteilen oberhalb des theoretischen Mittelpunkts der Skala noch etwas stärker ausgeprägt ($M=3,71$; $SD=0,85$). Sowohl die Leistungserwartungen als auch die Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens unterschieden sich in Abhängigkeit von den Hintergrundmerkmalen der Lernenden (vgl. Tab. 1). So fielen beispielsweise die Leistungserwartungen für türkischstämmige Kinder sowohl im sprachlichen als auch im mathematischen Bereich geringer aus als für Kinder ohne Zuwanderungshintergrund. Ihre Motivation und ihr Arbeitsverhalten beurteilten die Lehrkräfte aber genauso hoch wie von Kindern ohne Zuwanderungshintergrund. Außerdem schätzten die Lehrkräfte alle berücksichtigten Merkmale der Lernenden umso günstiger ein, je höher der soziale Status der Familien der Kinder war. Ebenfalls sprachen sie Mädchen eine bessere Leistungsentwicklung im sprachlichen Bereich, eine höhere Motivation und ein günstigeres Arbeitsverhalten zu als Jungen, wohingegen sie für Jungen einen stärkeren Leistungszuwachs in Mathematik erwarteten als für Mädchen.

Mit Blick auf die Testleistungen der Schülerinnen und Schüler (vgl. Tab. 2) kann festgehalten werden, dass die Unterschiede in den Leistungserwartungen zum Teil objektiv beobachtbare Leistungsunterschiede zwischen den Gruppen widerspiegeln. Teilweise ließen sich jedoch keine entsprechenden Unterschiede in den Testleistungen der Kinder identifizieren. Die Selbsteinschätzungen der Motivation und des Arbeitsverhaltens der Lernenden unterschieden sich vergleichsweise wenig zwischen den Gruppen. Lediglich Kinder aus sozial bessergestellten Familien gaben tendenziell eine höhere Anstrengungsbereitschaft an als Kinder aus sozial schwächeren Familien. Außerdem fiel die selbsteingeschätzte Lernfreude der Mädchen etwas höher aus als die der Jungen.

Um zu prüfen, ob sich die Befunde zur Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens domänenspezifisch unterscheiden, wurden die fächerbezogenen Items jeweils auch getrennt analysiert (vgl. Tab. 1). Dabei zeigte sich, dass die positivere Einschätzung der Motivation und des Arbeitsverhaltens von Mädchen bezogen auf den Deutschunterricht (Mädchen: $M=3,87$, $SD=0,87$; Jungen: $M=3,53$, $SD=0,84$, $t(871)=-5,82$, $p<0,05$) stärker ausfiel als bezogen auf das Fach Mathematik (Mädchen: $M=3,79$, $SD=0,89$; Jungen: $M=3,61$, $SD=0,88$, $t(838)=-2,97$, $p<0,05$). Bezüglich des ethnischen und sozialen Hintergrunds der Kinder waren keine Unterschiede zwischen den Domänen erkennbar. Eine Übersicht aller Interkorrelationen der Analyse- und Kontrollvariablen ist in Tab. A1 im Online Anhang zu finden.

Tab. 1 Mittelwerte und Standardabweichungen der Leistungserwartungen und der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens für verschiedene Schülergruppen

	Leistungserwartungen		Lehrkräfteeinschätzung Motivation und Arbeitsverhalten		
	Sprache	Mathe	Gesamt	Deutsch	Mathe
<i>Gesamt</i>	3,39 (0,94)	3,43 (0,85)	3,71 (0,85)	3,69 (0,87)	3,70 (0,89)
<i>Ethnischer Hintergrund</i>					
Deutschland	3,49 ^{c,d} (0,92)	3,46 ^{g,h} (0,85)	3,73 (0,85)	3,73 (0,86)	3,72 (0,89)
Türkei	2,98 ^{c,e} (0,96)	3,16 ^{g,i} (0,89)	3,55 ^k (0,87)	3,52 ^m (0,93)	3,50 ^o (0,92)
SU/Osteuropa	3,61 ^{e,f} (0,86)	3,76 ^{h,i,j} (0,79)	3,91 ^{k,l} (0,79)	3,88 ^{m,n} (0,81)	3,89 ^o (0,86)
Andere Länder	3,15 ^{d,f} (0,91)	3,29 ^j (0,82)	3,61 ^l (0,84)	3,56 ⁿ (0,87)	3,62 (0,88)
<i>Sozialer Hintergrund^a</i>					
Niedrig	3,00 (0,95)	3,11 (0,85)	3,45 (0,79)	3,42 (0,82)	3,40 (0,82)
Mittel	3,42 (0,91)	3,44 (0,83)	3,73 (0,87)	3,71 (0,89)	3,73 (0,91)
Hoch	3,75 (0,82)	3,75 (0,78)	3,95 (0,78)	3,97 (0,80)	3,96 (0,83)
<i>Geschlecht^b</i>					
Mädchen	3,48 (0,96)	3,34 (0,86)	3,88 (0,84)	3,87 (0,87)	3,79 (0,89)
Junge	3,31 (0,91)	3,51 (0,85)	3,55 (0,82)	3,53 (0,84)	3,61 (0,88)

Games-Howell-Test für ungleiche Varianzen (ethnischer und sozialer Hintergrund) bzw. t-Test (Geschlecht)

^aAlle Mittelwertsunterschiede zwischen den sozialen Gruppen auf dem 5%-Niveau signifikant

^bAlle Mittelwertsunterschiede zwischen den Geschlechtern auf dem 5%-Niveau signifikant

^{c-o}Mittelwertsunterschied zwischen den jeweiligen Gruppen auf dem 5%-Niveau signifikant

5.2 Unterschiede in der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens und Zusammenhänge mit Selbsteinschätzungen der Kinder

Zunächst wurde die Annahme geprüft, dass Lehrkräfte die Motivation und das Arbeitsverhalten von Schülerinnen und Schülern in Abhängigkeit von deren Hintergrundmerkmalen unterschiedlich einschätzen (Hypothesen 1a–c; vgl. Tab. 3).

Unter Kontrolle des SES und Geschlechts schätzten Lehrkräfte Kinder mit osteuropäischem Hintergrund als motivierter ein und sprachen ihnen ein günstigeres Arbeitsverhalten zu als Kindern ohne Zuwanderungshintergrund ($\beta_y = 0,26$, $p < 0,05$). Unterschiede für türkischstämmige Kinder im Vergleich zu Kindern ohne Zuwanderungshintergrund zeigten sich hingegen nicht ($\beta_y = -0,06$, $p = 0,51$). Lehrkräfte schrieben jedoch Kindern eine umso höhere Motivation und ein umso günstigeres

Tab. 2 Mittelwerte und Standardabweichungen der Leistungen sowie der Selbsteinschätzungen der Motivation und des Arbeitsverhaltens verschiedener Schülergruppen

	Leistungen			Selbsteinschätzung Motivation und Arbeitsverhalten	
	Lautbew.	Lesen	Mathe	Lernfreude	Anstreng.
<i>Gesamt</i>	13,54 (4,57)	23,99 (26,99)	27,58 (7,73)	1,54 (0,38)	1,70 (0,28)
<i>Ethnischer Hintergrund</i>					
Deutschland	13,77 ^a (4,56)	23,45 (25,97)	28,35 ^{d,e} (7,52)	1,52 (0,39)	1,70 (0,28)
Türkei	11,83 ^{a,b,c} (4,30)	19,53 (22,28)	24,70 ^{d,f} (7,74)	1,56 (0,35)	1,65 (0,29)
SU/Osteuropa	14,36 ^b (4,45)	29,98 (32,51)	29,07 ^{f,g} (7,54)	1,54 (0,39)	1,73 (0,26)
Andere Länder	13,35 ^c (4,65)	25,47 (29,49)	25,78 ^{e,g} (7,91)	1,58 (0,32)	1,72 (0,25)
<i>Sozialer Hintergrund</i>					
Niedrig	11,70 ^h (4,32)	17,82 ^{i,j} (19,71)	24,78 ^k (8,60)	1,51 (0,37)	1,64 ^{l,m} (0,31)
Mittel	13,67 ^h (4,37)	24,20 ⁱ (27,28)	27,90 ^k (7,10)	1,54 (0,38)	1,71 ^l (0,27)
Hoch	15,36 ^h (4,55)	30,19 ^j (30,78)	30,06 ^k (7,27)	1,57 (0,37)	1,76 ^m (0,23)
<i>Geschlecht</i>					
Mädchen	13,60 (4,51)	23,89 (24,40)	25,61 ⁿ (6,87)	1,59 ^o (0,34)	1,71 (0,26)
Junge	13,48 (4,64)	24,08 (29,13)	29,35 ⁿ (8,04)	1,50 ^o (0,40)	1,70 (0,29)

Games-Howell-Test für ungleiche Varianzen (ethnischer und sozialer Hintergrund) bzw. t-Test (Geschlecht)

^{a-o}Mittelwertsunterschied zwischen den jeweiligen Gruppen auf dem 5%-Niveau signifikant

Arbeitsverhalten zu, je höher der SES ihrer Familien war ($\beta_y = 0,22$, $p < 0,05$). Zudem wurden Mädchen im Vergleich zu Jungen im Durchschnitt als motivierter und anstrengungsbereiter eingeschätzt ($\beta_y = 0,42$, $p < 0,05$). Diese Unterschiede spiegeln sich nur zu einem sehr geringen Teil in der selbsteingeschätzten Lernfreude und Anstrengungsbereitschaft der Kinder wider und blieben auch nach Kontrolle dieser Merkmale mit geringfügig reduzierter Effektstärke bestehen (vgl. Tab. 3, Modell 2).

5.3 Vermittlung der Verzerrungen in Leistungserwartungen durch Lehrkrifteinschätzungen der Motivation und des Arbeitsverhaltens

Als Basis für die Mediationsanalysen wurden zunächst die Verzerrungen in den Leistungserwartungen der Lehrkräfte bestimmt. Neben den Hintergrundmerkmalen der Kinder wurden hierzu die bereichsspezifischen Ausgangsleistungen der Lernenden, ihre kognitiven Grundfähigkeiten sowie ihre selbsteingeschätzte Lernfreude und An-

Tab. 3 Prädiktion der Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens

	Modell 1			Modell 2		
	<i>b</i>	<i>SE b</i>	β_y	<i>b</i>	<i>SE b</i>	β_y
<i>Ethnischer Hintergrund (Referenz = Deutschland)</i>						
Türkei	-0,05	0,07	-0,06	-0,05	0,07	-0,06
SU/Osteuropa	0,22	0,08	0,26*	0,19	0,08	0,23*
Andere Länder	-0,05	0,06	-0,06	-0,09	0,07	-0,10
<i>Sozialer Hintergrund</i>						
Sozioökonomischer Status	0,19	0,03	0,22*	0,16	0,03	0,19*
<i>Geschlecht</i>						
Weiblich	0,35	0,06	0,42*	0,32	0,05	0,37*
<i>Selbsteinschätzung</i>						
Lernfreude	-	-	-	0,12	0,05	0,14*
Anstrengungsbereitschaft	-	-	-	0,14	0,04	0,16*
<i>Intercept</i>	3,53*	0,05	-	3,55*	0,05	-
<i>R</i> ²	10,2%*			17,2%*		

* $p < 0,05$

strebungsbereitschaft als Prädiktoren der Leistungserwartungen einbezogen. Trotz der leicht modifizierten Operationalisierung der Leistungserwartungen sind die gefundenen Verzerrungen mit den von Lorenz et al. (2016) berichteten Ergebnissen konsistent (vgl. Tab. 4 und 5, Modell 0). So erwarteten Lehrkräfte unter Kontrolle der Leistungen und selbsteingeschätzten Lernfreude und Anstrengungsbereitschaft der Kinder einen geringeren Leistungszuwachs im sprachlichen Bereich für türkischstämmige Kinder im Vergleich zu Kindern ohne Zuwanderungshintergrund ($\beta_y = -0,32$, $p < 0,05$), für Kinder aus sozial schwächeren Familien im Vergleich zu Kindern aus sozial bessergestellten Familien ($\beta_y = 0,15$, $p < 0,05$) und für Jungen im Vergleich zu Mädchen ($\beta_y = 0,17$, $p < 0,05$). Im mathematischen Bereich fielen die Leistungserwartungen höher aus für Kinder mit osteuropäischem Hintergrund als für Kinder ohne Zuwanderungshintergrund ($\beta_y = 0,34$, $p < 0,05$) und höher für Kinder aus sozial bessergestellten Familien als für Kinder aus sozial benachteiligten Familien ($\beta_y = 0,13$, $p < 0,05$). Geschlechtsbezogene Verzerrungen waren im mathematischen Bereich nicht zu verzeichnen ($\beta_y = -0,07$, $p = 0,34$).

Um die Hypothesen 2a und 2b zu prüfen, wurden die Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens als Mediator zwischen den Hintergrundmerkmalen der Kinder und den Leistungserwartungen hinzugenommen und indirekte Pfade spezifiziert (vgl. Tab. 4 und 5, Modell 1; siehe auch Abb. 1). Die Ergebnisse zeigen, dass je positiver die Lehrkräfte die Motivation und das Arbeitsverhalten der Kinder einschätzten, desto höher fielen insgesamt auch die Leistungserwartungen im sprachlichen und mathematischen Bereich aus (Sprache: $\beta_y = 0,52$, $p < 0,05$; Mathematik: $\beta_y = 0,54$, $p < 0,05$). Dieser Zusammenhang ergab sich bei gleichen Leistungen, bei gleicher selbsteingeschätzter Lernfreude und Anstrengungsbereitschaft sowie bei gleichen Hintergrundmerkmalen der Kinder.

Die Koeffizienten des *ethnischen Hintergrunds* in der Prädiktion der Leistungserwartungen änderten sich nach Kontrolle der Lehrkräfteeinschätzung der Motivation

Tab. 4 Prädiktion der sprachlichen Leistungserwartungen von Lehrkräften

	Modell 0			Modell 1			
	<i>b</i>	<i>SE b</i>	β_y	<i>b</i>	<i>SE b</i>	β_y	$\Delta\beta_y$
<i>Ethnischer Hintergrund (Referenz = Deutschland)</i>							
Türkei	-0,30	0,09	-0,32*	-0,28	0,08	-0,31*	-0,01
SU/Osteuropa	0,07	0,08	0,07	-0,01	0,06	-0,01	0,08
Andere Länder	-0,27	0,09	-0,29*	-0,23	0,08	-0,25*	-0,04
<i>Sozialer Hintergrund</i>							
SES	0,14	0,03	0,15*	0,07	0,03	0,08*	0,07*
<i>Geschlecht</i>							
Weiblich	0,15	0,07	0,17*	-0,02	0,05	-0,02	0,19*
<i>Leistungen</i>							
Lautbewusstheit	0,20	0,04	0,22*	0,13	0,03	0,14*	-
Lesen	0,20	0,03	0,22*	0,16	0,03	0,18*	-
<i>Kognitive Grundfähigkeiten</i>							
Kurzzeitgedächtnis	0,09	0,03	0,10*	0,03	0,02	0,03	-
Matrizentest	0,13	0,03	0,14*	0,07	0,03	0,08*	-
<i>Selbsteinschätzung</i>							
Lernfreude	0,07	0,04	0,08*	0,01	0,03	0,02	-
Anstrengungsbereitschaft	0,07	0,03	0,07*	0,02	0,03	0,02	-
<i>Lehrkräfteeinschätzung</i>							
Motivation/Arbeitsverhalten	-	-	-	0,49	0,04	0,52*	-
<i>Indirekte Effekte über Lehrkräfteeinschätzung Motivation/Arbeitsverhalten</i>							
Türkei	-	-	-	0,00	0,04	0,00	-
SU/Osteuropa	-	-	-	0,09	0,04	0,10*	-
Andere Länder	-	-	-	-0,03	0,04	-0,03	-
SES	-	-	-	0,10	0,02	0,11*	-
Weiblich	-	-	-	0,19	0,03	0,20*	-
<i>Intercept</i>	3,38*	-	-	3,47*	-	-	-
<i>R</i> ²	36,2%*	-	-	56,4%*	-	-	-

$\Delta\beta_y$ = Differenz von β_y zwischen Modell 0 und Modell 1

SES Sozioökonomischer Status

* $p < 0,05$

und des Arbeitsverhaltens weder im sprachlichen noch im mathematischen Bereich (alle $\Delta\beta_y$ nicht signifikant). Demnach scheinen die mit dem ethnischen Hintergrund verbundenen Verzerrungen der Leistungserwartungen unabhängig von der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens zu sein.

Anders stellte sich das Bild für Verzerrungen der Leistungserwartungen in Abhängigkeit vom *sozialen Hintergrund* der Kinder dar. Die Koeffizienten des SES reduzierten sich nach Kontrolle der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens signifikant (Sprache: $\Delta\beta_y = 0,07$, $p < 0,05$; Mathematik: $\Delta\beta_y = 0,08$, $p < 0,05$). Etwa die Hälfte der mit dem SES assoziierten Verzerrungen ließ sich auf Unterschiede in der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens der Kinder zurückführen. Die über diese Lehrkräfteeinschätzung vermittelten

Tab. 5 Prädiktion der mathematischen Leistungserwartungen von Lehrkräften

	Modell 0			Modell 1			
	<i>b</i>	<i>SE b</i>	β_y	<i>b</i>	<i>SE b</i>	β_y	$\Delta\beta_y$
<i>Ethnischer Hintergrund (Referenz = Deutschland)</i>							
Türkei	-0,06	0,06	-0,07	-0,07	0,05	-0,09	0,02
SU/Osteuropa	0,29	0,09	0,34*	0,21	0,06	0,25*	0,09
Andere Länder	0,00	0,08	0,00	0,01	0,06	0,01	-0,01
<i>Sozialer Hintergrund</i>							
SES	0,11	0,03	0,13*	0,05	0,02	0,05*	0,08*
<i>Geschlecht</i>							
Weiblich	-0,06	0,06	-0,07	-0,28	0,05	-0,33*	0,26*
<i>Leistungen</i>							
Mathematik	0,33	0,03	0,39*	0,21	0,03	0,25*	-
<i>Kognitive Grundfähigkeiten</i>							
Kurzzeitgedächtnis	0,04	0,02	0,05*	0,00	0,02	0,00	-
Matrizentest	0,13	0,03	0,15*	0,09	0,03	0,11*	-
<i>Selbsteinschätzung</i>							
Lernfreude	0,08	0,04	0,09*	0,02	0,03	0,03	-
Anstrengungsbereitschaft	0,06	0,04	0,07	0,02	0,03	0,02	-
<i>Lehrkräfteeinschätzung</i>							
Motivation/Arbeitsverhalten	-	-	-	0,45	0,03	0,54*	-
<i>Indirekte Effekte über Lehrkräfteeinschätzung Motivation/Arbeitsverhalten</i>							
Türkei	-	-	-	0,02	0,04	0,03	-
SU/Osteuropa	-	-	-	0,08	0,04	0,10	-
Andere Länder	-	-	-	-0,00	0,04	-0,00	-
SES	-	-	-	0,09	0,02	0,10*	-
Weiblich	-	-	-	0,18	0,03	0,21*	-
<i>Intercept</i>	3,43*	0,04	-	3,54*	0,04	-	-
<i>R</i> ²	36,7%*			57,7%*			

$\Delta\beta_y$ = Differenz von β_y zwischen Modell 0 und Modell 1

SES Sozioökonomischer Status

* $p < 0,05$

indirekten Pfade des SES waren statistisch signifikant (Sprache: $\beta_y = 0,11$, $p < 0,05$; Mathematik: $\beta_y = 0,10$, $p < 0,05$).

Für die *geschlechtsbezogenen* Verzerrungen in den Leistungserwartungen kehrten sich die Effekte um: Ohne Berücksichtigung der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens zeigte sich ein Vorteil zugunsten der Mädchen im sprachlichen Bereich und kein Geschlechtseffekt im mathematischen Bereich (vgl. Tab. 4 und 5, Modell 0). Wurde die Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens zusätzlich kontrolliert, verschwand der Geschlechtseffekt im sprachlichen Bereich ($\beta_y = -0,02$, $p = 0,67$) und im mathematischen Bereich trat ein Vorteil zugunsten der Jungen auf ($\beta_y = -0,33$, $p < 0,05$). Bei gleichen Testleistungen, gleicher selbsteingeschätzter Lernfreude und Anstrengungsbereitschaft der Kinder sowie bei gleicher Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens erwarteten die Lehrkräfte also für Jungen bessere Mathematikleistungen als

für Mädchen. Im sprachlichen Bereich erwarteten Lehrkräfte hingegen bei gleicher Leistung und bei gleicher Selbst- und Fremdeinschätzung der Motivation und des Arbeitsverhaltens ähnliche Leistungen für Mädchen und Jungen. Die über die Lehrkräfteinschätzungen der Motivation und des Arbeitsverhaltens vermittelten indirekten Pfade des Geschlechts waren statistisch signifikant (Sprache: $\beta_y = 0,20$, $p < 0,05$; Mathematik: $\beta_y = 0,21$, $p < 0,05$).

Wurden die fächerspezifischen Lehrkräfteinschätzungen der Motivation und des Arbeitsverhaltens in den Modellen berücksichtigt, veränderten sich die Ergebnisse im sprachlichen Bereich nicht bedeutsam. Auch im mathematischen Bereich blieb die Richtung der Befunde identisch. Allerdings fiel der Vorteil der Jungen nach Kontrolle der fächerspezifischen Lehrkräfteinschätzungen der Motivation und des Arbeitsverhaltens etwas geringer aus ($\beta_y = -0,23$, $p < 0,05$), wodurch auch der indirekte Pfad eine etwas geringere Stärke zeigte ($\beta_y = 0,10$, $p < 0,05$).

6 Diskussion

Der vorliegende Beitrag ging zwei Fragestellungen nach: Erstens, ob Lehrkräfteinschätzungen der Motivation und des Arbeitsverhaltens primär die Selbsteinschätzungen von Schülerinnen und Schülern oder eher stereotypenbasierte Annahmen widerspiegeln, und zweitens, ob ethnische, soziale und geschlechtsbezogene Verzerrungen von Leistungserwartungen auf diese Unterschiede zurückgeführt werden können.

Zunächst wurde deutlich, dass Lehrkräfte die Motivation und das Arbeitsverhalten von Schülerinnen und Schülern in Abhängigkeit von deren Hintergrundmerkmalen unterschiedlich einschätzten. Diese Differenzen spiegelten sich dabei nur marginal in den Selbsteinschätzungen der Kinder wider. Es kann also davon ausgegangen werden, dass es sich teilweise um stereotype Annahmen der Lehrkräfte handelt.

Übereinstimmend mit Befunden bisheriger Studien (z. B. Trautwein und Baeriswyl 2007; Anders et al. 2010) schrieben Lehrkräfte Kindern aus sozial bessergestellten Familien im Vergleich zu Kindern aus sozial schwächeren Familien sowie Mädchen im Vergleich zu Jungen eine höhere Motivation und ein günstigeres Arbeitsverhalten zu. Aus der Berücksichtigung der Selbsteinschätzungen der Kinder in der vorliegenden Studie ergibt sich darüber hinaus ein zusätzlicher Erkenntnisgewinn: Die Befunde verdeutlichen, dass die Einschätzungen der Lehrkräfte gegenüber Kindern aus sozial bessergestellten Familien und Mädchen auch dann noch höher ausfielen, wenn die Selbsteinschätzungen der Kinder berücksichtigt wurden. Die Ergebnisse bestätigen damit die Hypothesen 1b und 1c.

Unterschiede in Abhängigkeit vom ethnischen Hintergrund (Hypothese 1a) fanden sich unter Kontrolle des SES und Geschlechts nur für die Gruppe der Kinder mit osteuropäischem Hintergrund: Lehrkräfte schrieben diesen Schülerinnen und Schülern eine höhere Motivation und ein günstigeres Arbeitsverhalten zu als Kindern ohne Zuwanderungshintergrund, auch dann wenn die Selbsteinschätzungen der Kinder kontrolliert wurden. Dieser Befund steht im Gegensatz zu Studien von Anders et al. (2010) und Schneider (2011), die kleine negative Effekte des allgemeinen Zuwanderungshintergrunds auf die Lehrkräfteinschätzung der Motivation und

des Arbeitsverhaltens identifizierten. Allerdings wurden dort verschiedene ethnische Herkunftsgruppen nicht separat betrachtet, sodass sich entgegengesetzte Verzerrungen für verschiedene Zuwanderergruppen möglicherweise überdeckt haben.

Weiterhin zeigte sich, dass Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens mit höheren Leistungserwartungen assoziiert sind, wobei dieser Zusammenhang in Übereinstimmung mit Hypothese 2a auch unter Kontrolle der Testleistungen und der Selbsteinschätzungen der Kinder bestehen blieb. In die Leistungserwartungen von Lehrkräften flossen also nicht nur die tatsächlichen Leistungen der Lernenden sowie deren bei sich selbst wahrgenommene Lernfreude und Anstrengungsbereitschaft ein, sondern zusätzlich auch Lehrkraftannahmen zur Motivation und zum Arbeitsverhalten der Kinder.

Außerdem bestätigte sich die Annahme, dass Unterschiede in der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens in Abhängigkeit vom ethnischen und sozialen Hintergrund sowie vom Geschlecht der Schülerinnen und Schüler zur Erklärung von Verzerrungen in Leistungserwartungen beitragen (Hypothese 2b). Damit unterstützen die Befunde insgesamt die Vermutung, dass Verzerrungen in Leistungserwartungen von Lehrkräften teilweise auf Annahmen über die Motivation und das Arbeitsverhalten verschiedener Gruppen von Schülerinnen und Schülern basieren.

Die mit dem *sozialen Hintergrund* assoziierten Verzerrungen in Leistungserwartungen konnten etwa zur Hälfte auf Unterschiede in der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens zurückgeführt werden. Lehrkräfte scheinen anzunehmen, dass Kinder aus sozial bessergestellten Familien insgesamt motivierter und anstrengungsbereiter sind und dass sich damit einhergehend auch ihre Leistungen günstiger entwickeln. Der verbleibende verzerrte Anteil könnte z. B. auf die Annahme der Lehrkräfte zurückgehen, Kinder aus sozial bessergestellten Familien erhielten in schulischen Belangen mehr Unterstützung von ihren Eltern (vgl. z. B. Nölle et al. 2009). Zukünftige Studien sollten dieser Annahme nachgehen. Zusätzlich zu Lehrkräfteeinschätzungen der elterlichen Unterstützung sollten dabei auch Elternangaben berücksichtigt werden und als Vergleichsmaß für die Lehrkräfteeinschätzungen dienen.

Für die *geschlechtsbezogenen* Verzerrungen in den Leistungserwartungen kehrten sich die Effekte nach Kontrolle der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens um, sodass im mathematischen Bereich ein Vorteil zugunsten der Jungen und im sprachlichen Bereich kein Geschlechtseffekt auftrat. Demnach scheinen Lehrkräfte die Motivation und das Arbeitsverhalten von Jungen insgesamt als ungünstiger anzunehmen als von Mädchen und ihnen daher im sprachlichen Bereich einen Leistungsnachteil zu attestieren, obwohl sich die sprachlichen Leistungen der Mädchen und Jungen in unserer Studie nicht unterschieden. Im mathematischen Bereich scheinen die Lehrkräfte hingegen die höheren Leistungen, die sie den Jungen eigentlich zutrauen, nach unten zu adjustieren, weil sie Jungen als weniger schulisches motiviert einschätzen und ihnen ein ungünstigeres Arbeitsverhalten zusprechen. Zwar berichteten die Jungen auch tatsächlich eine etwas geringere Lernfreude als die Mädchen, diese Unterschiede wurden von den Lehrkräften jedoch überschätzt. Diese Befunde zu Leistungserwartungen und Geschlecht sind weitgehend konsis-

tent mit den Ergebnissen der bereits beschriebenen Studie von Kuhl und Hannover (2012) zur Benotung von Mädchen und Jungen.

Die mit dem *ethnischen Hintergrund* verbundenen Verzerrungen der Leistungserwartungen traten unabhängig von der Lehrkräfteeinschätzung der Motivation und des Arbeitsverhaltens auf. Dies weist darauf hin, dass Leistungserwartungen von Lehrkräften für Kinder aus zugewanderten Familien mit stereotypen Annahmen über andere lernrelevante Merkmale der Kinder oder ihres familiären Umfelds verknüpft sind. Ein möglicher Erklärungsansatz für die höheren Leistungserwartungen an Kinder mit osteuropäischem Hintergrund im mathematischen Bereich könnte sein, dass Lehrkräfte bei den Eltern dieser Kinder eine besondere Wertschätzung der mathematisch-naturwissenschaftlichen Fächer wahrnehmen (z. B. Walter und Taskinen 2008), was sich möglicherweise auch in höheren mathematischen Leistungserwartungen für diese Kinder niederschlagen könnte. Außerdem könnten geringere Erwartungen im sprachlichen Bereich für türkischstämmige Kinder durch die Annahme geringerer Deutschkenntnisse bedingt sein. So thematisierten z. B. in einer Interviewstudie von Kratzmann und Pohlmann-Rother (2012) alle befragten Erzieherinnen fehlende Deutschkenntnisse türkischstämmiger Kinder, die ihrer Einschätzung nach zu Herausforderungen im Kindergartenalltag führen. Allerdings sollte diesbezüglich bedacht werden, dass in unseren Analysen im sprachlichen Bereich die Lautbewusstheit als Leistungsmaß kontrolliert wurde. Da die Testung auf Deutsch erfolgte, bildet dieses Maß zu einem gewissen Anteil auch die sprachlichen Fähigkeiten der Kinder im Deutschen ab. Dennoch wäre es aufschlussreich, in weiterführenden Studien zusätzlich eine Gruppierung anhand des Sprachhintergrunds der Kinder vorzunehmen um eventuelle Unterschiede der Befunde im Vergleich zu den Ergebnissen zu ethnischen Herkunftsgruppen identifizieren zu können. Zukünftige Studien sollten außerdem klären, auf welche Merkmale sich stereotype Annahmen von Lehrkräften für verschiedene ethnische Herkunftsgruppen neben der Motivation und dem Arbeitsverhalten noch beziehen und ob diese die in der vorliegenden Studie nicht erklärten Anteile ethnischer Verzerrungen in Leistungserwartungen erklären können.

Die Ergebnisse des Vergleichs zwischen Selbst- und Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens legen nahe, dass es sich eher um stereotype Annahmen der Lehrkräfte über diese Merkmale der Lernenden handelt. Dennoch ist es auch möglich, dass die Lehrkräfteeinschätzungen tatsächliche Gruppenunterschiede im Lernverhalten der Kinder widerspiegeln, die die Erstklässlerinnen und Erstklässler selbst nicht beurteilen können. Zukünftige Studien sollten daher systematische Beobachtungen des Lernverhaltens der Kinder durchführen und diese als objektiveren Vergleichsmaßstab für die Lehrkräfteeinschätzungen heranziehen.

Bei der Interpretation der Ergebnisse insgesamt ist außerdem zu beachten, dass die Stichprobe der Schülerinnen und Schüler in Bezug auf ihre soziale Zusammensetzung leicht positiv selektiert war. Zudem haben sich möglicherweise besonders engagierte Lehrkräfte freiwillig für die Studienteilnahme entschieden. Geht man davon aus, dass engagiertere Lehrkräfte besonders bemüht sind, ihre Schülerinnen und Schüler präzise einzuschätzen, hätte die höhere Beteiligungsquote dieser Lehrkräfte an der Studie eher eine Unterschätzung der Verzerrungen in den Leistungserwartungen zur Folge. Der ermittelte Zusammenhang zwischen den Leistungserwartungen

und den Lehrkräfteeinschätzungen der Motivation und des Arbeitsverhaltens sollte davon jedoch nicht beeinflusst sein.

Werden nicht erklärte Varianzanteile in Leistungserwartungen als Verzerrungen interpretiert, ist es möglich, dass Messungenauigkeiten in den Leistungsmaßen zu einer Unterschätzung des zutreffenden Anteils der Leistungserwartungen und damit zu einer Überschätzung etwaiger Verzerrungen führen. Trotz des Einbezugs vielfältiger Leistungsmaße lässt sich dies in der vorliegenden Studie nicht ausschließen.

Insgesamt weisen die Befunde darauf hin, wie wichtig es ist, Lehrkräfte für den möglichen Einfluss von Stereotypen auf schulische Einschätzungen und Erwartungen zu sensibilisieren. Dies ist vor allem deshalb relevant, weil verzerrte Leistungserwartungen über Prozesse der selbsterfüllenden Prophezeiung die Leistungsentwicklung Heranwachsender beeinflussen (Jussim und Harber 2005) und damit zu einer Verfestigung von Bildungsungleichheiten beitragen können. Dass es möglich ist, den Einfluss von Stereotypen zu reduzieren, zeigt die sozialpsychologische Literatur (vgl. z. B. Blair 2002). Hierzu ist es jedoch erforderlich, die zugrunde liegenden Mechanismen zu kennen und diesen aktiv entgegen zu wirken. Hierzu hat die vorliegende Studie wichtige Erkenntnisse geliefert.

Literatur

- Anders, Y., McElvany, N., & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte? In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule: Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 313–330). Bonn: Bundesministerium für Bildung und Forschung, Referat Bildungsforschung.
- Anderson-Clark, T.N., Green, R.J., & Henley, T.B. (2008). The relationship between first names and teacher expectations for achievement motivation. *Journal of Language and Social Psychology*, 27(1), 94–99.
- Artelt, C., Blossfeld, H.-P., Faust, G., Roßbach, H.-G., & Weinert, S. (2013). Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS-3-10) (Version: 2. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz). https://doi.org/10.5159/IQB_BIKS_3_10_v2.
- Bäuerlein, K., Beinicke, A., Berger, N., Faust, G., Jost, M., & Schneider, W. (2012). *FIPS Fähigkeitsindikatoren Primarschule: Ein computerbasiertes Diagnoseinstrument zur Erfassung der Lernausgangslage und der Lernentwicklung von Schulanfängern. Manual*. Göttingen: Hogrefe.
- Baumert, J., Watermann, R., & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell. *Zeitschrift für Erziehungswissenschaft*, 6(1), 46–71.
- van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R.W. (2010). The implicit prejudiced attitudes of teachers: relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, 47(2), 497–527.
- Blair, I.V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6(3), 242–261.
- Böhmer, I., Hörstermann, T., Gräsel, C., Krolak-Schwerdt, S., & Glock, S. (2015). Eine Analyse der Informationssuche bei der Erstellung der Übergangsempfehlung: Welcher Urteilsregel folgen Lehrkräfte? *Journal für Bildungsforschung Online*, 7(2), 59–81.
- Diehl, C., & Fick, P. (2016). Ethnische Diskriminierung im deutschen Bildungssystem. In C. Diehl, C. Hunkler, & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf* (S. 243–286). Wiesbaden: Springer VS.
- 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(1), 198–208.

- 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. In M. P. Zanna (Hrsg.), *Advances in experimental social psychology* (Bd. 23, S. 1–74). San Diego: Academic Press.
- Ganter, S. (1997). *Stereotype und Vorurteile: Konzeptualisierung, Operationalisierung und Messung*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung.
- Ganzeboom, H. B. G. (2010). *A new international socio-economic index (ISEI) of occupational status for the international standard classification of occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007*. Annual Conference of International Social Survey Programme, Lisbon.
- Givvin, K. B., Stipek, D. J., Salmon, J. M., & MacGyvers, V. L. (2001). In the eyes of the beholder: students' and teachers' judgments of students' motivation. *Teaching and Teacher Education*, 17(3), 321–331.
- Hachfeld, A., Anders, Y., Schroeder, S., Stanat, P., & Kunter, M. (2010). Does immigration background matter? How teachers' predictions of students' performance relate to student background. *International Journal of Educational Research*, 49(2–3), 78–91.
- Heinze, A., Herwartz-Emden, L., & Reiss, K. (2007). Mathematikkenntnisse und sprachliche Kompetenz bei Kindern mit Migrationshintergrund zu Beginn der Grundschulzeit. *Zeitschrift für Pädagogik*, 53(4), 562–581.
- Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology*, 101(3), 662–670.
- Hirschauer, M., & Kullmann, H. (2010). Lehrerprofessionalität im Zeichen von Heterogenität – Stereotype bei Lehrkräften als kollegial zu bearbeitende Herausforderung. In J. Hagedorn, V. Schurt, C. Steber & W. Waburg (Hrsg.), *Ethnizität, Geschlecht, Familie und Schule* (S. 351–373). Wiesbaden: Verlag für Sozialwissenschaften.
- Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: a review of literature. *Review of Educational Research*, 59(3), 297–313.
- Jussim, L., & Eccles, J. (1992). Teacher expectations II: construction and reflection of student achievement. *Journal of Personality and Social Psychology*, 63(6), 947–961.
- Jussim, L., & Harber, K. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9(2), 131–155.
- Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: accuracy and the quest for the powerful self-fulfilling prophecy. In M. P. Zanna (Hrsg.), *Advances in experimental social psychology* (Bd. 28, S. 281–388). San Diego: Academic Press.
- Karing, C. (2009). Diagnostische Kompetenz von Grundschul- und Gymnasiallehrkräften im Leistungsbereich und im Bereich Interessen. *Zeitschrift für Pädagogische Psychologie*, 23(3–4), 197–209.
- Karing, C., Matthäi, J., & Artelt, C. (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I – Eine Frage der Spezifität? *Zeitschrift für Pädagogische Psychologie*, 25(3), 159–172.
- Kratzmann, J., & Pohlmann-Rother, S. (2012). Ethnische Stereotype im Kindergarten? Erzieherinnenhaltungen gegenüber Zuwanderern aus der Türkei. *Zeitschrift für Pädagogik*, 58(6), 855–876.
- Kuhl, P., & Hannover, B. (2012). Differenzielle Benotungen von Mädchen und Jungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), 153–162.
- Lorenz, G. (2017). *Selbsterfüllende Prophezeiungen in der Schule: Leistungserwartungen von Lehrkräften und Kompetenzen von Kindern mit Zuwanderungshintergrund*. Wiesbaden: Springer VS.
- 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.
- McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, 46(3), 235–261.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Nölle, I., Hörstermann, T., Krolak-Schwerdt, S., & Gräsel, C. (2009). Relevante diagnostische Informationen bei der Übergangsempfehlung – die Perspektive der Lehrkräfte. *Unterrichtswissenschaft*, 37(4), 294–310.
- Prenzel, M., Sälzer, C., Klieme, E., & Köller, O. (Hrsg.) (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Rauer, W., & Schuck, K. D. (2004). *FEESSI-2: Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen*. Göttingen: Beltz.

- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: the role of child background and classroom context. *American Educational Research Journal*, 48(2), 335–360.
- Richter, D., Kuhl, P., & Pant, H. A. (2012). Soziale Disparitäten des Kompetenzerwerbs. 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. 191–207). Münster: Waxmann.
- Rjosk, C., McElvany, N., Anders, Y., & Becker, M. (2011). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung der basalen Lesefähigkeit ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht*, 58(2), 92–105.
- Schneider, T. (2011). Die Bedeutung der sozialen Herkunft und des Migrationshintergrundes für Lehrerurteile am Beispiel der Grundschulempfehlung. *Zeitschrift für Erziehungswissenschaft*, 14(3), 371–396.
- Schrader, F.-W. (2010). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 102–108). Weinheim: Beltz.
- Schrader, F.-W., & Helmke, A. (1990). Lassen sich Lehrer bei der Leistungsbeurteilung von sachfremden Gesichtspunkten leiten? Eine Untersuchung zu Determinanten diagnostischer Lehrerurteile. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 22(4), 312–324.
- Schrader, F.-W., & Helmke, A. (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 45–58). Weinheim: Beltz.
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19(1–2), 85–95.
- Spinath, B. (2011). Lernmotivation. In H. Reinders, H. Ditton, C. Gräsel, & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 45–55). Wiesbaden: Springer VS.
- Stahl, N. (2007). Schülerwahrnehmung und -beurteilung durch Lehrkräfte. In H. Ditton (Hrsg.), *Kompetenzaufbau und Laufbahnen im Schulsystem* (S. 171–198). Münster: Waxmann.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: a meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762.
- Trautwein, U., & Baeriswyl, F. (2007). Wenn leistungsstarke Klassenkameraden ein Nachteil sind: Referenzgruppeneffekte bei Übertrittsentscheidungen. *Zeitschrift für Pädagogische Psychologie*, 21(2), 119–133.
- Walter, O., & Taskinen, P. (2008). Naturwissenschaftsbezogene Motivationen und Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund in Deutschland. Der Einfluss der Generationen, der Herkunft und des Elternhauses. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (S. 185–203). Wiesbaden: Verlag für Sozialwissenschaften.
- Weiß, R. H., & Osterland, J. (1997). *Grundintelligenztest Skala 1: CFT 1. Handanweisung* (5. Aufl.). Göttingen: Hogrefe.