Original Article

Teacher Stereotypes and Teacher Expectations at the Intersection of Student Gender and Socioeconomic Status

Sarah Gentrup[1], Melanie Olczyk[2], and Georg Lorenz[3,4]

[1]Department of Education Studies, Humboldt-Universität zu Berlin, Germany
[2]Institute of Sociology, Martin-Luther-Universität Halle-Wittenberg, Germany
[3]Education Department, University of Potsdam, Germany
[4]Department of Sociology, Leipzig University, Germany

Abstract: Gender and family socioeconomic status (SES) are central dimensions of educational inequality and may interact in shaping inequality. This study addresses teacher expectations and stereotypes that possibly contribute to intersectional inequality. The study relies on two samples of teachers and students in German primary schools (sample 1: $N_{Teachers}$=69, 94% female, $N_{Students}$= 1,049 (German language)/ 1,027 (mathematics), 48% female; sample 2: $N_{Teachers}$= 698 (German language)/614 (mathematics), 94% female; $N_{Students}$= 4,732 (German language)/4,117 (mathematics), 51% female). Two-level regression analyses revealed additive gender and socioeconomic bias in teacher expectations in German language and mathematics but no intersectional bias (i.e., constant gender bias along family SES and similar socioeconomic bias for girls and boys). Further, teachers with more traditional gender stereotypes showed stronger gender bias, while SES-related stereotypes were unrelated to teacher expectations. We discuss how additive teacher expectation biases relate to the broader concept of intersectionality, potentially shaping unique educational experiences at the intersection of gender and SES.

Keywords: educational disparities, intersectionality, stereotypes, teacher expectations

Stereotype und Erwartungen von Lehrkräften an der Schnittstelle von Geschlecht und sozioökonomischem Status der Schüler_innen

Zusammenfassung: Geschlecht und sozioökonomischer Hintergrund sind zentrale Dimensionen bildungsbezogener Ungleichheiten, die Bildungsungleichheiten wechselseitig prägen können. Die vorliegende Studie untersucht Erwartungen und Stereotype von Lehrkräften, die zu intersektionalen Bildungsungleichheiten beitragen können. Die Studie basiert auf zwei Stichproben von Lehrkräften und Schüler_innen in Grundschulen in Deutschland (Stichprobe 1: $N_{Lehrkräfte}$=69, 94% weiblich, $N_{Schüler\_innen}$=1,049 (Deutsch)/1,027 (Mathematik), 48% weiblich; Stichprobe 2: $N_{Lehrkräfte}$=698 (Deutsch)/614 (Mathematik), 94% weiblich; $N_{Schüler\_innen}$=4,732 (Deutsch)/4,117 (Mathematik), 51% weiblich). Zwei-Ebenen-Regressionsanalysen zeigten additive Verzerrungen der sprachlichen und mathematischen Leistungserwartungen in Abhängigkeit vom Geschlecht und sozioökonomischen Hintergrund, jedoch keine intersektionalen Verzerrungen (d.h., geschlechtsbezogene Verzerrungen unabhängig vom sozioökonomischen Hintergrund und vergleichbare Verzerrungen in Abhängigkeit vom sozioökonomischen Hintergrund für Mädchen und Jungen). Zudem zeigten Lehrkräfte mit traditionelleren Geschlechterstereotypen stärker nach dem Geschlecht verzerrte Leistungserwartungen, während sozioökonomische Stereotype nicht bedeutsam mit Verzerrungen zusammenhingen. Wir diskutieren, inwiefern additive Erwartungsverzerrungen mit der generellen Idee von Intersektionalität übereinstimmen, da sie spezifische Bildungserfahrungen an der Schnittstelle von Geschlecht und sozioökonomischem Status generieren können.

Schlüsselwörter: Bildungsungleichheit, Intersektionalität, Lehrkrafterwartungen, Stereotype

Intersectional theoretical frameworks have received growing attention in multiple research areas, including educational research (e.g., Parker et al., 2020). These frameworks recognize that social categories, such as gender, ethnicity, and socioeconomic status (SES), are not only separately related to the distribution of power in society but also in their simultaneity and interdependency (Cole, 2009; Else-Quest & Hyde, 2016). By applying a framework of intersectionality, educational researchers acknowledge that each student simultaneously shares multiple social categories that jointly shape educational experiences and interact in forming inequality (Else-Quest & Hyde, 2016). Because of the roots of intersectionality in Black feminism (Crenshaw, 1989, 1991), inter-

sectional research in psychology has predominantly focused on gender in combination with race or ethnicity, while paying less attention to the intersection of gender with other characteristics, such as SES (Eagly et al., 2012).

Students' gender and family SES are two essential sources of educational inequality (for recent overviews, Keller et al., 2022; Liu et al., 2022; Rosén et al., 2022). Girls have been found, on average, to outperform boys in language skills, such as reading (Mullis et al., 2023; recently, for primary school students in Germany, Gentrup et al., 2022), whereas boys often have been observed to perform slightly better than girls in mathematics and science (Else-Quest et al., 2010; Mullis et al., 2020; recently, for primary school students in Germany, Gentrup et al., 2022). Also, students from lower-SES families were disadvantaged in academic achievement compared to students from higher-SES families (Harwell et al., 2017; Liu et al., 2022; recently, for primary school students in Germany, Sachse et al., 2022). However, existing research also stresses that gender and socioeconomic inequality are interdependent (e.g., Parker et al., 2020). Students' family SES proved to be more closely related to boys' school achievement than to girls', putting boys from lower-SES families in an especially disadvantaged situation (Entwisle et al., 2007; Lühe et al., 2017; Mensah & Kiernan, 2010; Penner & Paret, 2008; for evidence on cross-country variation of these effects, Brunner et al., 2023; Eriksson & Lindvall, 2023). In mathematics, this pattern resulted in more pronounced gender gaps (to the advantage of boys) among students from higher-SES families than among students from lower-SES families (Lühe et al., 2017; Mensah & Kiernan, 2010; Penner & Paret, 2008). In contrast, the gender gap to the advantage of girls in the language domain was less pronounced among students from higher-SES families than among students from lower-SES families (Entwisle et al., 2007; Lühe et al., 2017; Mensah & Kiernan, 2010).

One factor possibly contributing to such (intersectional) educational disparities are teachers and their expectations for their students' capability. When teacher expectations inaccurately vary for different groups of students (i.e., are biased), such effects can contribute to educational disparities (e.g., Lorenz, 2021; Olczyk, Gentrup et al., 2023; Olczyk, Kwon et al., 2023). While gender bias and socioeconomic bias in teacher expectations have widely been shown in educational research (for an overview, Wang et al., 2018), it remains unclear whether and how student gender and family SES interrelate in forming teacher expectations. Also, to understand a possible interplay of gender and family SES in shaping teacher expectations, we need further knowledge about how stereotypes are sources of bias in teacher expectations (Fiske & Neuberg, 1990). While some evidence has illustrated the effects of teacher stereotypes on bias in teacher expectations (Glock et al., 2016; Lorenz, 2021; Lorenz et al., 2023; Muntoni & Retelsdorf, 2018; Tiedemann, 2002), it lacks an intersectional perspective that considers a possible interplay of multiple stereotypes.

First, our study addresses this gap by replicating earlier work by examining student gender and family SES and their effects on teacher expectation bias. Second, we extend current understanding by investigating (a) how these two categories interact in shaping teacher expectation biases and (b) which teacher stereotypes influence the formation of teacher expectations for girls and boys from different SES backgrounds. To ensure the robustness of our results, we rely on two datasets targeting primary education in Germany, each allowing for a comparison of the language domain – a female-stereotyped domain – and mathematics – a male-stereotyped domain.

Theory and Evidence

Evidence on Gender and Socioeconomic Bias in Teacher Expectations

Teacher expectations are inferences teachers make about their students' current and future achievements (e.g., Ready & Chu, 2015). Even though they mirror the actual test scores of students to a substantial degree (Hoge & Coladarci, 1989; Südkamp et al., 2012), teacher expectations have also proved to be biased regarding ascribed student characteristics, such as gender, family SES, or ethnic origin (for a recent overview, Wang et al., 2018).

Regarding *gender bias* in teacher expectations, empirical research yielded domain-specific results (Wang et al., 2018): Teachers tended to overestimate girls' skills in language, such as reading or literacy (e.g., Lorenz et al., 2016; Meissel et al., 2017; Ready & Wright, 2011), and boys' skills in mathematics (e.g., Gentrup & Rjosk, 2018; Tiedemann, 2002). Regarding *socioeconomic bias*, the research reports that teachers often overestimate the academic skills of students from more advantaged families and underestimate those from disadvantaged families (Wang et al., 2018). Such bias is observed across various skills and domains, including reading and math achievement (e.g., Lorenz et al., 2016; Olczyk, Kwon et al., 2023; Ready & Wright, 2011).

Many studies simultaneously investigated two or more student characteristics and their relationship to teacher expectations (e.g., Lorenz et al., 2016; van den Bergh et al., 2010). This approach covers the additive effects of student characteristics on teacher expectations. Studies relying on an intersectional theoretical approach are

scarce and focus mostly on the intersection of ethnicity and gender (Bonefeld et al., 2022; Shepherd, 2011). Research is particularly limited regarding the intersection of students' gender and family SES. Auwarter and Aruguete (2008) explored teacher judgments of personal characteristics (e.g., competence) and future expectations (e.g., dropping out of high school) using vignettes of fictional students, who demonstrated academic and behavioral problems. The results revealed that teachers assumed more favorable characteristics for boys from higher-SES families than lower-SES families but less favorable ones for girls from higher-SES families. Thus, family SES contributed differently to teachers' perceptions of girls and boys with academic and behavioral problems. However, in the same study, teachers also expected a less promising future for students from lower-SES families regardless of the students' gender, which does not point to interdependent effects. In light of this scarce and mixed evidence on intersectionality in teacher expectation bias, we continue to refer to social-psychological models to draw considerations on how student gender and family SES may interact in forming teacher expectations.

Stereotypes as Sources of Bias in Teacher Expectations

Researchers explain bias in teacher expectations with stereotypes, which are cognitive structures that contain generalized beliefs about the characteristics of members of social groups (e.g., generalized beliefs about the math achievement of girls; Eagly, 1987). According to dual-process models, like the *continuum model* of impression formation (Fiske & Neuberg, 1990), human perception shifts between stereotype-based and attribute-based information-processing modes. The stereotype-based mode begins automatically, conserving mental resources by categorizing individuals into groups based, for example, on their gender or SES. Given sufficient time, cognitive resources, motivation, and available individual information, impression formation can shift toward attribute-based modes, where more detailed individual information becomes integrated into the perception, with gender and SES being just some of the many attributes. This mode results in a detailed picture of the target but requires the integration of a vast amount of information and, thus, time and cognitive resources that are often scarce.

Students' gender and family SES are crucial sources of social categorizations (Eagly, 1987; Lindqvist et al., 2017; Stolier & Freeman, 2016) and shape teacher expectations (e.g., Wang et al., 2018). Studies have shown that teachers assume girls like learning, put effort into school, and are tidy and self-disciplined (Duckworth & Seligman, 2006; Gentrup et al., 2018; Glock & Kleen, 2017). Domain-specific stereotypes include the belief that mathematics and science are male domains, whereas the language domain is a female domain (Heyder & Kessels, 2013; Steffens & Jelenec, 2011). Regarding SES, teachers often assume students from higher-SES families are more focused and motivated in school, receive more parental support, are more intelligent, and, thus, perform better throughout all school domains than students from lower-SES families (Schuchardt & Dunkake, 2014; Tobisch & Dresel, 2020).

Few studies have assessed teacher stereotypes and investigated whether they account for bias in teachers' expectations. Muntoni and Retelsdorf (2018) showed that fifth-grade teachers who more strongly endorsed a reading-female stereotype expected higher reading achievement for girls than teachers who shared a more gender-egalitarian view about reading. Based on data from third and fourth grade, Tiedemann (2002) found similar results indicating a stronger expectation bias in favor of boys in mathematics for teachers more strongly endorsing a mathematics-male stereotype. Regarding students' family SES, in their vignette study, Glock and colleagues (2016) identified a stronger positive bias in the achievement expectations toward students from higher-SES families for teachers who had more positive implicit and explicit attitudes toward students from higher-SES families. However, these studies do not allow for conclusions on whether and how gender stereotypes and SES-related stereotypes interact in creating bias in teacher expectations.

How Stereotype Activation and Expectation Bias May Work in the Case of Multiple Student Characteristics

It is an open research question of how teachers deal – cognitively – with multiple options for categorization: Which stereotype(s) become(s) activated? Does stereotype activation result in intersectional patterns of teacher expectation bias? Social-psychological models make varying predictions regarding these questions. As Petsko and Bodenhausen (2020) outlined, there are three major theoretical perspectives regarding stereotyping in the case of multiple person characteristics: (1) the integration perspective, (2) the dominance perspective, and (3) the compartmentalization perspective.

The *integration perspective* assumes that a perceiver integrates all social categories a person belongs to in their perception (e.g., Freeman & Ambady, 2014; Freeman et al., 2020; Hall et al., 2019). Based on this view, teacher

expectations of student characteristics should be colored by gender stereotypes and SES-related stereotypes as well as by an infinite number of further social categories to which a student belongs. This process would result in an almost attribute-based impression formation (Fiske & Neuberg, 1990) in which stereotypes do not become very powerful (Petsko & Bodenhausen, 2020). As a result, teacher expectations should be largely unbiased by the students' gender and SES.

The *dominance perspective*, in contrast, suggests that some social categories are dominant over others, unavoidably grabbing the perceivers' attention and suppressing the perception of other social categories (e.g., Sidanius et al., 2018). Gender and age are social categories seen as specifically powerful (Glick & Fiske, 1996; Sidanius et al., 2018; Stangor et al., 1992). This perspective, therefore, would lead to the assumption that student gender dominates automatic ways of teacher expectation formation, resulting in gender bias but no socioeconomic bias.

The *compartmentalization perspective* lies between these two perspectives (Petsko & Bodenhausen, 2020). It assumes that social contexts influence which social category receives attention and gains mental dominance (e.g., the *activation-inhibition model* by Bodenhausen & Macrae, 1998; for details on relevant situational factors, see Petsko & Bodenhausen, 2020). Based on this perspective, either the gender or the SES-related stereotype should influence a teacher's perception of a student. However, recent work on the compartmentalization perspective also recognizes the possibility of genuine intersectional stereotypes (e.g., about girls from higher-SES families; *lens-based model of intersectional stereotyping* by Petsko et al., 2022). Based on this view, the one "lens" (Petsko et al., 2022) through which a teacher evaluates the student can be based on one social category or an intersection of two (or more) social categories. That is, the teacher may evaluate a female student from a lower-SES family as a girl, as a student from a lower-SES family, or specifically as a girl from a lower-SES family (i.e., not as a girl in general, not as lower-SES in general, but specifically as a lower-SES girl).

The Current Study

Our study expands the state of research on bias in teacher expectations in reading and mathematics by considering the interdependency of gender and family SES. Moreover, we examine which teacher stereotypes come into play during the formation of the teachers' expectations for girls and boys from higher- and lower-SES families. Our study also supplements experimental work by providing evidence from school surveys. We investigate the following two research questions: (1) What is the extent of bias in teacher expectations at the intersection of gender and family SES? (2) Can teachers' gender and SES-related stereotypes account for gender and socioeconomic bias in teacher expectations?

Based on the three perspectives of stereotyping in the case of multiple student characteristics mentioned above, one could assume different patterns of bias and stereotype application. As the *integration perspective* postulates that an infinite number of social categories a student belongs to would shape teacher expectations, we should expect an almost attribute-based perception process resulting in largely unbiased expectations. Based on the *dominance perspective*, student gender should be the relevant category that contributes to teacher expectation bias, because, according to social-psychological research, gender is the more salient social category and should thus shape impression formation to a greater extent than family SES (e.g., Sidanius et al., 2018; Stangor et al., 1992). This would manifest itself in gender bias but not socioeconomic bias. The *compartmentalization perspective* would suggest that, in specific situations, either gender *or* family SES *or* their intersection would be dominant during teachers' impression formation. Thus, in specific situations, only gender bias *or* socioeconomic bias *or* intersectional bias should occur. However, because we investigate survey data that display patterns of bias across situations, these processes would manifest in either additive or interdependency biases within the data.

Our study explores different scenarios of teacher expectation bias. Specifically, we investigate the following patterns of bias: (1) no bias, (2) gender bias but no socioeconomic bias, (3) socioeconomic bias but no gender bias, (4) additive gender and socioeconomic bias but no interdependent bias, (5) and interdependent bias (i.e., gender bias varying between students from lower- and higher-SES families and/or socioeconomic bias varying between girls and boys). Further, we relate the observed bias(es) in teacher expectations to teachers' gender and SES-related stereotypes.

To ensure the robustness of our results, we chose to analyze two datasets targeting the beginning of primary education in Germany, each allowing for a comparison of the language domain – a female-stereotyped domain – and mathematics – a male-stereotyped domain. One dataset stems from a study specifically designed to assess teacher expectations and possible bias reducing the risk that student achievement already adapted to teacher expectations, but was conducted in only one German federal state (for details, see Gentrup et al., 2020). The second dataset stems from a large-scale assessment representatively

covering all German federal states (for details, see Blossfeld & Roßbach, 2019) and, thus, allowing for more generalizable conclusions.

Methods

Databases

We analyzed data from two data sources from the project "Kompetenzerwerb und Lernvoraussetzungen" (KuL; English translation: Competence Acquisition and Learning Preconditions) and the German National Educational Panel Study (NEPS; Blossfeld & Roßbach, 2019).[1]

The KuL study was specifically designed to investigate teacher expectations and their effects on the students' development throughout first grade. Teacher-student contact was minimal prior to the measurement of teacher expectations and student learning-related characteristics, reducing the likelihood of previous self-fulfilling prophecies (for detailed information, Gentrup et al., 2020). The KuL study was conducted during the school year 2013/2014 in primary schools in North Rhine-Westphalia, Germany. The analytic sample contained information from 69 teachers (94 % female; average age 42 years, *SD* = 9.16) as well as from 1,049 first-grade students in 65 classes (German language) and from 1,027 first-grade students in 64 classes (mathematics). On average, the students were 6 years old (*SD* = 0.33) upon entering school.

The NEPS is an ongoing national longitudinal study. In the present contribution, we used data from the Starting Cohort 2 (NEPS-SC2), which is assumed to represent children in Grade 1 in the school year 2012/2013. Specifically, we examined data from Wave 4, Grade 2, which was mainly assessed in the last quarter of 2013 and was the earliest grade level in which teachers' stereotypes were collected. We considered only cases where the teacher who rated individual students actually instructed the respective school subject (i. e., cases in which the German language teacher had rated students' skills in German language and in which the mathematics teachers had rated students' mathematical skills, respectively). The analyzed sample contained information from 4,732 second-grade students and 698 teachers (German language) and from 4,117 second-grade students and 614 teachers (mathematics). The teachers (94 % female) were on average 46 years old (*SD* = 10.73). The students were on average 8 years old (*SD* = 0.39).

Instruments

Teacher Expectations

A few weeks after school enrollment, the teachers participating in the KuL study reported their expectations for each student's skills in German language (e. g., reading and writing words) and mathematics (e. g., handling numbers and quantities) for the upcoming year. They compared these skills to the class average using a 5-point scale, where 1 indicates *much worse* and 5 indicates *much better*. The items were aggregated in a mean score, separately for German language (α = 0.94) and mathematics (α = 0.94, Gentrup et al., 2020).

In the NEPS, teachers rated students' written language abilities (i. e., ability to understand and write texts) and mathematical skills (i. e., dealing with numbers and quantities) in the middle of Grade 2, relative to the skills of same-aged children, on a 5-point-scale (1 = *much worse* to 5 = *much better*).

Teacher Stereotypes

In KuL and NEPS, the teachers were asked, "What results do you think [first (KuL)/second (NEPS)] graders from the following groups achieve overall in the competence field 'reading' compared to [first (KuL)/second (NEPS)] graders in Germany?" The same question was asked for mathematics. The teachers reported their beliefs regarding the average reading and math achievement of different student groups (i. e., female students, male students, students from higher-SES families, and students from lower-SES families) on a Likert scale ranging from 1 (*far below average*) to 11 (*far above average*) in KuL and from 1 (*far below average*) to 10 (*far above average*) in NEPS (Wenz et al., 2016). To cover the extent of stereotypes, we calculated the difference between teacher evaluations for female and male students (i. e., for German language: female minus male; for mathematics: male minus female) as well as for students from higher- and lower-SES families in both domains (i. e., higher-SES minus lower-SES). Higher scores, thus, indicate that a teacher assumes a greater advantage for girls in German language, for boys in mathematics, and for students from higher-SES families in both domains, respectively.

Student Gender

We included student gender in our analyses as a dichotomized variable (the coding was varied in different models; see the section "Analytic Strategy" for further information). We generated the variable based on information from multiple sources and different time points: In

[1] The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.

KuL, we used student and parent reports; in the NEPS, parent reports and institutional student lists provided information on the students' gender. In KuL, 48% of the students were female; in NEPS, the share of female students was 51%.

Student Family SES

Family SES was measured using the *Highest International Socioeconomic Index of Occupational Status* (HISEI; Ganzeboom et al., 1992) among the students' parents, who were asked to complete a parental questionnaire. The HISEI can vary between the values of 11 and 89. The mean values in the two datasets were M_{KuL} = 52.69 (SD_{KuL} = 19.60) and M_{NEPS} = 56.73 (SD_{NEPS} = 19.87).

Student Immigrant Background

Because earlier studies identified ethnic bias in teacher expectations (e.g., Lorenz et al., 2016) and family SES and immigrant status overlap substantially, we controlled for the immigrant status using a dummy-coded variable (0 = *no immigrant status*, 1 = *foreign-born student and/or both parents foreign-born*). We generated the variable based on parental reports of their own and their child's country of birth. In KuL and NEPS, 25% and 11% of the students had an immigrant background, respectively.

Students' Learning-Related Characteristics

To identify bias in teacher expectations, we considered students' learning-related characteristics, such as achievement, as accurate predictors of teacher expectations (see section "Identification of Teacher Expectation Bias"). In KuL, the German version of the Performance Indicators in Primary Schools (PIPS; Bäuerlein et al., 2012) served to objectively assess students' language abilities (phonological awareness: α = .82; reading: α = .96) and mathematical skills (α = .92) at the beginning of the school year. Further, students completed a deductive reasoning test (Culture Fair Intelligence Test [CFT]; Weiß & Osterland, 1997; α = .78) and the subscale *Working memory* implemented in the PIPS assessment (α = .76), both capturing students' general cognitive abilities. In the middle of the school year, the students also indicated their motivation, i.e., their enjoyment of learning (α = .78) and the effort they invest in learning (α = .70) on a 3-point scale ranging from 0 (*not true*) to 2 (*completely true*; measured with an adapted version of the FEESS1–2; Rauer & Schuck, 2004). We calculated sum scores from all achievement test scores and used the students' motivation as mean scores.

In the NEPS, we considered the following achievement tests: For students' German language skills, we used the information on the students' early reading skills measured in Grade 2 with the modified ELFE test (sum score; Lenhard & Schneider, 2006), the students' receptive vocabulary tested in Grade 1 (measured by a modified Peabody Picture Vocabulary Test [PPVT]; Berendes et al., 2013; WLE reliability = .87, Fischer & Durda, 2020), and the students' receptive grammar skills tested in Grade 1 (measured by the Test for Reception of Grammar [TROG]; Lorenz et al., 2017; WLE reliability = .77, Welling & Zink, 2023). For the mathematical skills, we used results from achievement tests constructed by the NEPS and assessed in Grade 2 covering content-related (e.g., quantity, space, and shape) and process-related components (e.g., applying technical skills, modeling, problem-solving; WLE reliability = .79, Schnittjer & Gerken, 2018). Except for early reading skills, we used IRT-scaled scores (i.e., WLEs) for each of these measures, as provided in the Scientific Use Files. As an indicator of nonverbal cognitive abilities, we used results from the NEPS-MAT test (sum score; α = .69, Lang et al., 2014) administered at Grade 2.

Analytic Strategy

Identification of Teacher Expectation Bias

To identify any inaccuracy in teacher expectations, we followed a residual approach that was suggested by Madon and colleagues (1997) and applied in other studies (e.g., Gentrup et al., 2020). To this end, we conducted multiple regression analyses with cluster robust standard errors to predict teacher expectations by the students' learning-related characteristics (see Electronic Supplementary Material ESM 1). We then saved the residuals of these regressions as they contained the portion of the teacher expectations not predicted by the students' actual characteristics and can be interpreted as teacher expectation bias. As teacher expectations cover predictions about students' future achievements, we decided to include not only domain-specific achievement as accurate sources of the teacher expectations into the regressions but also further student characteristics that are known to determine learning processes, such as general cognitive abilities and motivation. We thus applied a more conservative approach that reduces the risk of overestimating teacher expectation bias.

A positive residual score indicated that a teacher had more positive expectations for a student's achievement than was predicted by that student's learning-related characteristics (i.e., teacher overestimation); a negative score reflected more negative expectations for a student (i.e., teacher underestimation). Values close to zero represented an accurate prediction of student achievement and, hence, indicate unbiased expectations (Madon et al., 1997). In the second step, these residuals were used as

Table 1. Descriptive statistics by study

	M/%	*SD*	Min.	Max.
KuL				
Teacher gender stereotypes				
German language	1.10	1.38	−2.00	4.00
Mathematics	0.75	0.98	−1.00	3.00
Teacher SES-related stereotypes				
German language	5.00	1.99	1.00	10.00
Mathematics	3.07	2.05	0.00	8.00
Teacher expectation bias				
German language	−0.01	0.79	−2.46	2.12
Mathematics	0.00	0.69	−2.48	2.63
HISEI	52.69	19.60	14.21	88.96
Female students	48%		0	1
Students of immigrant descent	25%		0	1
NEPS				
Teacher gender stereotypes				
German language	1.26	1.29	−2.00	7.00
Mathematics	0.14	1.38	-7.00	4.00
Teacher SES-related stereotypes				
German language	4.49	2.14	−2.00	10.00
Mathematics	3.60	2.15	−6.00	10.00
Teacher expectation bias				
German language	0.00	0.89	−3.75	2.84
Mathematics	0.00	0.88	−2.75	2.89
HISEI	56.73	19.87	11.56	88.96
Female students	51%		0	1
Students of immigrant descent	11%		0	1

Notes. N by construct and study: teacher expectation bias: KuL: $N_{students}$ = 1,049 (German language) and $N_{students}$ = 1,027 (mathematics); NEPS: $N_{students}$ = 4,732 (German language) and $N_{students}$ = 4,117 (mathematics); teacher stereotypes: KuL: $N_{teachers}$ = 69; NEPS: $N_{teachers}$ = 698 (German language) and $N_{teachers}$ = 614 (mathematics); student background characteristics: KuL: $N_{students}$ = 1,049; NEPS: $N_{students}$ = 4,958. All descriptive statistics were calculated prior to *z*-standardization and were based on the first imputed dataset. *Sources.* Own calculations based on KuL and NEPS-SC2.

dependent variables to test our research questions (see next section).

Investigation of Gender and Socioeconomic Bias in Teacher Expectations

We conducted multilevel models with random intercepts separately for both German language and mathematics using Stata (StataCorp, 1985–2023). These models consider that teacher expectations for each student (level 1) were nested in teachers (level 2).

All metric variables were *z*-standardized to allow for comparisons between the two samples. Descriptive statistics of the study variables are displayed in Table 1 (for further information on the distribution of the teacher stereotypes, see ESM 2).

To investigate gender and socioeconomic bias, we used the residuals from the first analysis step as a dependent variable (see previous section) and regressed them on the students' gender and family SES (Model 1). By adding an interaction term, we considered a potential interdependent bias along student gender and family SES (Model 2). We modeled all predictors in theses regression analyses at level 1 (student level).

To determine whether teacher expectation bias stemmed from teachers' gender and SES-related stereotypes, we added the teachers' gender and SES-related stereotypes as predictors at level 2 (teacher level) to our models. We first examined the cross-level interaction between student gender and gender stereotypes to see whether the effect of the teachers' gender stereotypes varied between girls and boys (Model 3). We then considered a three-way cross-level interaction between student gender, family SES, and gender stereotypes to investigate whether the effect of the teachers' gender stereotypes varied at the intersection of student gender and family SES (Model 4). Next, we analyzed whether the effect of the teachers' SES-

related stereotypes changed with the students' family SES by considering the cross-level interaction between family SES and SES-related stereotypes (Model 5). Finally, we explored the three-way cross-level interaction between student gender, family SES, and SES-related stereotypes to determine whether the effect of the teachers' SES-related stereotypes varied at the intersection of student gender and family SES (Model 6).[2]

To improve readability, we report gender-specific slope coefficients for the female and male students within the main text (Table 2). We calculated these coefficients by running all analyses twice: first, with male students forming the reference group (i. e., 0 = *male*, 1 = *female*) and, second, with female students forming the reference group (i. e., 0 = *female*, 1 = *male*). This approach helps ease interpreting three-way interactions (e. g., gender x HISEI x SES-related stereotypes). In addition to Table 2 in the main text, we present classic regression tables with the complete models, including coefficients for the controls and values for R^2 for both gender codings in ESM 3.

We used multiple imputations with iterated chained equations to include cases with missing information (MICE; White et al., 2011). We imputed missing data of all analytic variables (except student gender) separately for the student and the teacher samples. In KuL, the share of missing data varied between 2% and 15% at the student level and between 14% and 23% at the teacher level. In the NEPS, there were 1% to 10% missing data at the student level and 17% to 27% at the teacher level. Besides the variables used in the analyses, the imputation models also included further information from teacher and parent interviews as well as further results from student achievement tests (for detailed information on the imputation models and auxiliary variables, see the codes available on the Open Science Framework). We generated 50 imputed datasets. We conducted the regression analyses individually for each of the 50 imputed datasets and subsequently pooled their parameters according to Rubin's rules (Rubin, 1987). The imputations achieved convergence (StataCorp, 2023). Internal checks such as comparing summary statistics between the observed and the imputed data also substantiated the appropriateness of the selected imputation models (for an overview, see Nguyen et al., 2017).

Results

The complete results of the two-level multivariate regression analyses are displayed in ESM 3 (Tables E1 and E3 show the results when male students are the reference group and Tables E2 and E4 when female students are the reference group). The tables in ESM 3 also display coefficients for the controls as well as values for R^2. To ease the interpretation of the two- and three-way interactions in Models 2, 3, 4, and 6, we summarized the slope coefficients regarding family SES, stereotypes and their interactions for female and male students in Table 2.

Bias in Teacher Expectations

German Language

Regarding German language, Model 1 indicated a gender bias to the advantage of girls (KuL: β = .22, p < .001; NEPS: β = .21, p < .001) and a socioeconomic bias to the advantage of students from higher-SES families in both studies (KuL: β = .20, p < .001; NEPS: β = .16, p < .001). Model 2 additionally accounted for a possible interaction between student gender and family SES in predicting teacher expectations. For both studies, the interaction term was nonsignificant (KuL: β = −.05, p = .460; NEPS: β = .05, p = .064), indicating that the relationship between family SES and teacher expectations was comparable for girls and boys in KuL (girls: β = .17, p < .001; boys: β = .22, p < .001) and in NEPS (girls: β = .19, p < .001; boys: β = .13, p < .001).

Mathematics

In mathematics, the results of Model 1 indicated no gender bias in KuL (β = −.02, p = .755) but a significant gender bias to the advantage of boys in NEPS (β = −.27, p < .001). In both datasets, we observed a socioeconomic bias to the advantage of students from higher-SES families (KuL: β = .16, p < .001; NEPS: β = .10, p < .001). Similar to Model 2 in German language, the interaction term between student gender and family SES was nonsignificant in KuL (β = −.03, p = .692) and NEPS (β = .05, p = .097). Thus, the relationship between students' family SES and teacher expectations was comparable for girls and boys in KuL (girls: β = .14, p = .005; boys: β = .17, p < .001) and in NEPS (girls: β = .12, p < .001; boys: β = .07, p = .002).

[2] We reanalyzed all models using the highest parental education as an indicator of the students' family SES (instead of HISEI). Also, based on the KuL data, we re-ran models 3 and 4 using an alternative scale capturing teacher gender stereotypes. Further information on the robustness checks and their results are presented in ESM 4. Overall, the robustness checks lead to results comparable to the main results.

Table 2. Slope coefficients of students' gender and family SES as well as teachers' gender and SES-related stereotypes on teacher expectation bias in German language and mathematics for female and male students

		KuL						NEPS					
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Teacher expectation bias German language													
Female:		.22***	.22***	.22***	.22***	.22***	.22***	.21***	.21***	.21***	.21***	.21***	.21***
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Gender independent[a]:	HISEI	.20***						.16***					
		(0.03)						(0.02)					
Female:	HISEI		.17***	.17***	.17***	.17***	.16***		.19***	.18***	.18***	.18***	.19***
			(0.05)	(0.05)	(0.05)	(0.05)	(0.05)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Male:	HISEI		.22***	.22***	.22***	.22***	.22***		.13***	.13***	.13***	.13***	.13***
			(0.05)	(0.05)	(0.05)	(0.05)	(0.05)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Female:	Gender stereotype			−.04	−.04					.03	.03		
				(0.07)	(0.07)					(0.03)	(0.03)		
Male:	Gender stereotype			−.13	−.13					.00	.00		
				(0.07)	(0.07)					(0.03)	(0.03)		
Female:	HISEI x Gender stereotype				.01						−.01		
					(0.05)						(0.02)		
Male:	HISEI x Gender stereotype				−.02						.01		
					(0.05)						(0.02)		
Gender independent[a]:	SES-related stereotype					−.02						.01	
						(0.05)						(0.02)	
Female:	SES-related stereotype						−.02						−.00
							(0.06)						(0.03)
Male:	SES-related stereotype						−.02						.03
							(0.06)						(0.03)
Gender independent[a]:	HISEI x SES-related stereotype					.05						.03	
						(0.03)						(0.02)	
Female:	HISEI x SES-related stereotype						.09						.01
							(0.05)						(0.02)
Male:	HISEI x SES-related stereotype						.01						.04
							(0.04)						(0.02)
Teacher expectation bias mathematics													
Female:		-.02	-.02	-.02	-.02	-.01	-.01	-.27***	-.27***	-.27***	-.27***	-.27***	-.27***
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)

Table 2. Slope coefficients of students' gender and family SES as well as teachers' gender and SES-related stereotypes on teacher expectation bias in German language and mathematics for female and male students (Continued)

		KuL						NEPS					
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Gender independent[a]:	HISEI	.16***						.10***					
		(0.04)						(0.02)					
Female:	HISEI		.14**	.14**	.14**	.14**	.14**		.12***	.12***	.12***	.12***	.12***
			(0.05)	(0.05)	(0.05)	(0.05)	(0.05)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Male:	HISEI		.17***	.17***	.16***	.17***	.17***		.07**	.07**	.07**	.07**	.07**
			(0.05)	(0.05)	(0.05)	(0.05)	(0.05)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Female:	Gender stereotype			−.09	−.09					−.06[b*]	−.06[b*]		
				(0.06)	(0.06)					(0.03)	(0.03)		
Male:	Gender stereotype			−.00	−.00					.04[b]	.04[b]		
				(0.06)	(0.06)					(0.03)	(0.03)		
Female:	HISEI x Gender stereotype				−.05						−.01		
					(0.05)						(0.03)		
Male:	HISEI x Gender stereotype				.02						−.02		
					(0.05)						(0.02)		
Gender independent[a]:	SES-related stereotype					.13*						-.01	
						(0.05)						(0.02)	
Female:	SES-related stereotype						.14*						−.03
							(0.06)						(0.03)
Male:	SES-related stereotype						.12*						.02
							(0.06)						(0.03)
Gender independent[a]:	HISEI x SES-related stereotype					-.03						.00	
						(0.03)						(0.02)	
Female:	HISEI x SES-related stereotype						−.02						.01
							(0.05)						(0.03)
Male:	HISEI x SES-related stereotype						−.03						−.00
							(0.05)						(0.02)

Notes. Standardized regression coefficients. Standard errors in parentheses. SES = socioeconomic status. HISEI = Highest International Socio-Economic Index of Occupational Status. KuL: $N_{students}$ = 1,049 and $N_{teachers}$ = 65 (German language), $N_{students}$ = 1,027 and $N_{teachers}$ = 64 (mathematics); NEPS: $N_{students}$ = 4,732 and $N_{teachers}$ = 698 (German language), $N_{students}$ = 4,117 and $N_{teachers}$ = 614 (mathematics). * $p < .05$, ** $p < .01$, *** $p < .001$. [a] Models 1 and 5 did not include any interaction terms with the students' gender; therefore, the regression coefficients for the respective variables are gender-independent regression coefficients. [b] The difference between the regression coefficients for female and male students is statistically significant ($p < .05$). *Sources.* Own calculations based on KuL and NEPS-SC2.

Teacher Stereotype Effects

German Language

We continued by evaluating whether teachers' gender and SES-related stereotypes contribute to the observed gender and socioeconomic bias in teacher expectations. The results are displayed in Models 3 to 6 (see Table 2 for gender-specific slopes and ESM 3 for the complete regression models).

Regarding German language, the analyses indicated no substantial relationships between the teachers' gender or SES-related stereotypes and gender or socioeconomic bias in teacher expectations. This applied equally to both datasets for girls and boys from families with different SES.

Mathematics

The results were somewhat different for mathematics. Regarding gender bias, Models 3 and 4 revealed a significant interaction between gender and gender stereotypes in the NEPS data, indicating that teachers with stronger gender stereotypes underestimated girls more strongly than did teachers with more egalitarian beliefs ($\beta = -.06$, $p = .019$). For boys, the teachers' gender stereotypes were unrelated to the teachers' expectations ($\beta = .04$, $p = .159$). The difference between the two stereotype-related coefficients was statistically significant ($\beta = -.10$, $p = .001$). Based on the KuL data, we observed no relationships between gender stereotypes and teacher expectations.

SES-related stereotypes significantly predicted teacher expectations in Model 5 based on the KuL data ($\beta = .13$, $p = .006$). Not confirming an interdependency of gender- and SES-related bias in teacher expectations means this effect was independent of students' gender and family SES (all interaction terms were nonsignificant). Thus, teachers who more strongly endorsed SES-related stereotypes had inaccurately higher mathematical expectations for all of their students (i.e., they overestimated all of their students more strongly) than teachers with less pronounced SES-related stereotypes. This indicates that teachers who assume stronger achievement-related advantages of students from higher-SES families have higher expectations for their students' scholastic skills. However, SES-related stereotypes did not predict teacher expectation bias in the KuL data when we used the parents' educational level instead of HISEI (see ESM 4). Therefore, this result is not robust against using an alternative indicator of students' family SES. In the NEPS data, we found no significant effects of SES-related stereotypes or their interactions.

Discussion

The present study addressed teacher stereotypes related to gender and family SES as sources of a possible additive or interdependent teacher expectation bias. First, we studied the extent of bias in teacher expectations at the intersection of gender and family SES. Second, we examined whether teachers' gender stereotypes and SES-related stereotypes account for gender and socioeconomic bias in teacher expectations.

After controlling for student background variables, including domain-specific achievements, cognitive abilities, and, in KuL, students' motivation, we showed that teachers have higher achievement expectations for girls in German language and for boys in mathematics. We also found evidence that points to a socioeconomic bias that is additive to gender bias: After controlling for student gender and background variables, we found that teachers had higher expectations for students from higher-SES families than for students from lower-SES families in both German language and mathematics. These results indicate that teacher expectation biases along gender and family SES can exist independently and additively. These results agree with the compartmentalization perspective (Petsko & Bodenhausen, 2020), which predicts that dominant social categories can influence a person's perception independently or even interdependently. Our findings also mirror the results of a broad range of studies on teacher expectation bias (for an overview, see Wang et al., 2018). However, socioeconomic bias in teacher expectations did not differ between girls and boys, and gender bias did not vary depending on the students' family SES. This finding applied to both school domains and to both datasets.

In addition to gathering evidence on teacher expectation bias, our study also investigated more global stereotypes toward genders and students of different SES backgrounds and their relationships with teacher expectation bias. We found that teachers who have more pronounced gender stereotypes (i.e., believe that girls perform better in German language while boys perform better in mathematics) showed stronger gender bias in their achievement expectations, at least in mathematics and in the NEPS sample. Teachers who believe that boys perform better in mathematics than girls underestimate the mathematical skills of girls more strongly than teachers who believe that girls and boys perform equally in mathematics. This agrees with the basic notion that stereotypes shape social perception (Fiske & Neuberg, 1990) and with empirical findings by Tiedemann (2002). We could not confirm, though, that SES-related teacher stereotypes relate to socioeconomic bias in teacher ex-

pectations. Our findings, therefore, contradict evidence based on the student vignettes that supported stronger socioeconomic bias when teachers had more negative implicit and explicit attitudes toward students from lower-SES families (Glock et al., 2016). Importantly, we also found no evidence that teachers exhibit intersectional stereotypes, at least at the intersection of gender and family SES. In the samples we analyzed, gender stereotypes were not associated with socioeconomic bias in teacher expectations, and SES-related stereotypes were unrelated to gender bias in teacher expectations. Further, SES-related stereotypes were similarly associated with teacher expectations for girls and boys, while gender stereotypes related similarly to expectations for lower- and higher-SES students' skills in our data. This evidence agrees with the activation-inhibition model (Bodenhausen & Macrae, 1998) but not with the additional notion of intersectional stereotypes formulated in the lens-based model of intersectional stereotyping (Petsko et al., 2022). Our results suggest that both gender and SES independently influence how students are perceived. Thus, the impact on teacher expectations of having a lower-SES background does not differ significantly between girls and boys.

Note that the teacher stereotypes measured in this study may substantially reflect an accurate knowledge of actual gender and social disparities. This does not challenge our interpretation that teachers' group-specific performance beliefs are stereotypes but in fact highlights that stereotypes can be accurate (Jussim et al., 2018). Accurate stereotypes can lead to individual misestimations but not to systematic bias at the group level (Lorenz, 2021). This condition might (partly) explain why gender stereotypes did not predict teacher expectation bias in German language. Specifically, actual test score differences between the genders are more pronounced and were more consistently observed in language than in mathematics (e.g., Gentrup et al., 2022). Therefore, gender stereotypes predicting that "girls are better in reading than boys" are comparatively more accurate than gender stereotypes predicting that "boys are better in math than girls."

Limitations and Future Research

Our study is not without limitations. First, some findings vary between the two datasets under investigation, which might stem from the different sample sizes in KuL and NEPS and possible differences in statistical power. Also, the learning-related student characteristics we controlled for when identifying bias in teacher expectations partly varied between the datasets (e.g., self-reported student motivation was available in KuL only). An additional explanation may lie in the smaller variation of the teacher stereotypes in the teacher sample of the KuL study than in the NEPS study (for details, see ESM 2). These small differences in the endorsement of stereotypes might result from age differences in the teacher samples in KuL and NEPS. Also, participation in the KuL study was voluntary, and the participation rate was only 4%. Therefore, the teachers in KuL may have been positively selected. The stereotypes and achievement expectations among teachers in the KuL sample may thus have been less biased than among teachers in the NEPS sample. Another reason for differences between both samples might be that the answer category of the stereotype item in KuL contained a middle category (11-point scale), allowing (more) teachers to provide neutral answers to the stereotype questions in the KuL survey but not in the NEPS survey (10-point scale). But we should also note that the results of both datasets support the study's key findings and overall conclusions: In both samples, we found additive gender- and SES-related bias in the teacher expectations but no evidence for interdependencies between these biases.

Another limitation may be that our stereotype measures referred only to scholastic performance; stereotypes regarding other school-related characteristics, such as students' motivation, learning habits, or parental support (Gentrup, 2020; Gentrup et al., 2018), might be more predictive of the observed bias in teacher expectations. Future studies should explore these possibilities.

Finally, stereotypes could be intersectional (Petsko et al., 2022), in a manner that could not be tested with our data. First, this could be the case because the data we analyzed included measures of singular stereotypes, referring solely to gender or family SES; measures of genuine intersectional stereotypes (e.g., about girls from higher-SES families) were unavailable in both datasets. Second, what type of stereotype is activated among teachers (i.e., singular or intersectional) can depend on situational attributes, including the accessibility of intersectional stereotypes, stereotype fit, perceiver goal, and distinctiveness (Petsko & Bodenhausen, 2020). For instance, different students could fit differently into specific intersectional stereotypes, which may create variation in how singular stereotype measures (like the ones used in the current study) predict teacher expectations for girls and boys with different family SES. Similarly, the composition of a specific class could shape whether a singular or an intersectional stereotype contributes more information to teacher expectations. Future studies should measure intersectional stereotypes *and* situational factors to test these possibilities, preferably using experimental research designs. This would also allow for testing the different theoretical perspectives (i.e., the integration, dominance

and compartmentalization perspectives) in greater depth. Moreover, future studies should test whether intersectional stereotypes and interdependent biases among teachers exist at other intersections, such as along gender and special educational needs status or family SES and ethnic origin.

Conclusion

Research shows that gender inequality in scholastic performance differs with students' family SES (e.g., Lühe et al., 2017). According to our results, this pattern of interdependency does not likely stem from bias in teacher expectations, as we found additive gender and socioeconomic bias but no interdependency between these biases. Nevertheless, our results do align with the more general notion of intersectionality (e.g., Cole, 2009; Crenshaw, 1989, 1991; Else-Quest & Hyde, 2016) because additive bias suggests that teachers have different expectations for the achievement of girls with a lower family SES than for the achievement of boys with a lower family SES (as much as teacher expectations differ between boys with lower family SES and boys with a higher family SES). Specifically, the teacher expectations we identified in this study suggest the lowest reading competencies for boys from lower-SES families and higher reading competencies for girls from lower-SES families. Moreover, the results mirror the belief of lower mathematical skills among girls with a lower family SES than boys with a lower family SES. Thus, girls and boys from families with different SES experience different expectations from their teachers based on their individual combination of gender and SES. Should these teacher expectations translate into actual differences in educational performance – creating a self-fulfilling prophecy as suggested by Gentrup et al. (2020) – they would particularly disadvantage lower-SES girls in mathematics and lower-SES boys in German language.

Our finding that teachers' gender stereotypes in mathematics relate to gender bias in teacher expectations suggests that teachers' generalized beliefs can be a source of bias in their student evaluations. This highlights that teachers should be aware of such beliefs and their effects in everyday teaching. Although we found no evidence for the effects of SES-related stereotypes on teacher expectation bias, such stereotypes may still bias evaluations. Making teachers aware of stereotypes as generalized beliefs prevalent in everyone's mind appears unavoidable as part of teacher education. This applies particularly to the social category of gender. In line with our results, teachers should be in particular sensitized that the (implicit) assumption of mathematics as a "boys' domain" can have severe consequences for the evaluation and scholastic performance of girls and for gender inequality in the STEM area.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1026/0049-8637/a000291

ESM 1. Results of regression analyses calculating teacher expectation bias.

ESM 2. Distributions of the teachers' gender and SES-related stereotypes.

ESM 3. Results of the multilevel regression analyses.

ESM 4. Robustness checks.

References

Auwarter, A. E., & Aruguete, M. S. (2008). Effects of student gender and socioeconomic status on teacher perceptions. *Journal of Educational Research*, *101*(4), 243 – 246. Retrieved from http://www.jstor.org/stable/27548242

Bäuerlein, K., Beinicke, A., Berger, N., Faust, G., Jost, M., & Schneider, W. (2012). *Fähigkeitsindikatoren Primarschule (FIPS): Ein computerbasiertes Diagnoseinstrument zur Erfassung der Lernausgangslage und der Lernentwicklung von Schulanfängern* [Performance indicators in primary schools (PIPS): A computer-based diagnostic tool for detecting the initial learning capabilities and the learning development of children starting school]. Hogrefe.

Berendes, K., Weinert, S., Zimmermann, S., & Artelt, C. (2013). Assessing language indicators across the lifespan within the German national educational panel study (NEPS). *Journal for Educational Research Online, 5*(2), 15 – 49. https://doi.org/10.25656/01:8423

Blossfeld, H.-P., & Roßbach, H.-G. (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS).* Springer VS. https://doi.org/10.1007/978-3-658-23162-0

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer (Ed.), *Advances in social cognition* (Vol. 11, pp. 1 – 52). Erlbaum.

Bonefeld, M., Kleen, H., & Glock, S. (2022). The effect of the interplay of gender and ethnicity on teachers' judgements: Does the school subject matter? *The Journal of Experimental Education, 90*(4), 818 – 838. http://doi.org/10.1080/00220973.2021.1878991

Brunner, M., Keller, L., Stallasch, S. E., Kretschmann, J., Hasl, A., Preckel, F., Lüdtke, O., & Hedges, L. V. (2023). Meta-analyzing individual participant data from studies with complex survey designs: A tutorial on using the two-stage approach for data from educational large-scale assessments. *Research Synthesis Methods, 14*(1), 5 – 35. https://doi.org/10.1002/jrsm.1584

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64*(3) 170 – 180. https://doi.org/10.1037/a0014564

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago*

Legal Forum, 1989(1), 139 – 167. Retrieved from https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review, 43*, 1241 – 1299. https://doi.org/10.2307/1229039

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. https://doi.org/10.1037/0022-0663.98.1.198

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation.* Psychology Press. https://doi.org/10.4324/9780203781906

Eagly, A. H., Eaton, A., Rose, S. M., Riger, S., & McHugh, M. C. (2012). Feminism and psychology: Analysis of a half-century of research on women and gender. *American Psychologist, 67*(3), 211 – 230. https://doi.org/10.1037/a0027260

Else-Quest, N. M., & Hyde, J. S. (2016). Intersectionality in quantitative psychological research: I. Theoretical and epistemological issues. *Psychology of Women Quarterly, 40*(2), 155 – 170. https://doi.org/10.1177/0361684316629797

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103 – 127. https://doi.org/10.1037/a0018053

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education, 80*(2), 114 – 138. https://doi.org/10.1177/003804070708000202

Eriksson, K., & Lindvall, 743 J. (2023). Cultural variation in the SES-gender interaction in student achievement. *Frontiers in Psychology, 14*, Article 1120211. https://doi.org/10.3389/fpsyg.2023.1120211

Fischer, L., & Durda, T. (2020). *NEPS technical report for receptive vocabulary: Scaling results of Starting Cohort 2 for Kindergarten (Wave 1), Grade 1 (Wave 3) and Grade 3 (Wave 5)* (NEPS Survey Paper No. 65). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://doi.org/10.5157/NEPS:SP65:1.0

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 (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1 – 74). Academic Press.

Freeman, J. B., & Ambady, A. (2014). The dynamic interactive model of person construal coordinating sensory and social processes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 235 – 248). Guilford.

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Dynamic interactive theory as a domain-general account of social perception. *Advances in Experimental Social Psychology, 61*, 237 – 287. https://doi.org/10.1016/bs.aesp.2019.09.005

Ganzeboom, H. B. G., Graaf, P. M. D., & Treiman, D. J. (1992). A standard international socioeconomic index of occupational status. *Social Science Research, 21*, 1 – 56.

Gentrup, S., & Rjosk, C. (2018). Pygmalion and the gender gap: Do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling? *Educational Research and Evaluation, 24*(3 – 5), 295 – 323. https://doi.org/10.1080/13803611.2018.1550840

Gentrup, S., Rjosk, C., Stanat, P., & Lorenz, G. (2018). Einschätzungen der schulischen Motivation und des Arbeitsverhaltens durch Grundschullehrkräfte und deren Bedeutung für Verzerrungen in Leistungserwartungen [Teachers' perceptions of students' motivation and learning behavior and their role in biased teacher achievement expectations]. *Zeitschrift für Erziehungswissenschaft, 21*, 867 – 891. https://doi.org/10.1007/s11618-018-0806-2

Gentrup, S., Lorenz, G., Kristen, C., & Kogan, I. (2020). Self-fulfilling prophecies in the classroom: Teacher expectations, teacher feedback and student achievement. *Learning and Instruction, 66*, Article 101296. https://doi.org/10.1016/j.learninstruc.2019.101296

Gentrup, S., Schipolowski, S., & Wittig, J. (2022). Geschlechtsbezogene Disparitäten [Gender-based disparitites]. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 127 – 150). Waxmann. https://doi.org/10.31244/9783830996064

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491 – 512.

Glock, S., & Kleen, H. (2017). Gender and student misbehavior: Evidence from implicit and explicit measures. *Teaching and Teacher Education, 67*, 93 – 103. https://doi.org/10.1016/j.tate.2017.05.015

Glock, S., Krolak-Schwerdt, S., & Hörstermann, T. (2016). The higher the SES the better? Implicit and explicit attitudes influence preservice teachers' judgments of students. In C. Fields (Ed.), *Stereotypes and stereotyping: Misperceptions, perspectives and role of social media* (pp. 1 – 16). Nova Science Publishers.

Hall, E. V., Hall, A. V., Galinsky, A. D., & Phillips, K. W. (2019). MOSAIC: A model of stereotyping through associated and intersectional categories. *Academy of Management Review, 44*(3), 643 – 672. https://doi.org/10.5465/amr.2017.0109

Harwell, M., Maeda, Y., Bishop, K., & Xie A. (2017). The surprisingly modest relationship between SES and educational achievement. *The Journal of Experimental Education, 85*(2), 197 – 214. https://doi.org/10.1080/00220973.2015.1123668

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles, 69*(11 – 12), 605 – 617. https://doi.org/10.1007/s11199-013-0309-9

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research, 59*(3), 297 – 313. https://doi.org/10.3102/00346543059003297

Jussim, L., Stevens, S.T., & Honeycutt, N. (2018). Unasked questions about stereotype accuracy. *Archives of Scientific Psychology, 6*(1), 214 – 229. https://doi.org.10.1037/arc0000055

Keller, L., Preckel, F., Eccles, J. S., & Brunner, M. (2022). Top-performing math students in 82 countries: An integrative data analysis of gender differences in achievement, achievement profiles, and achievement motivation. *Journal of Educational Psychology, 114*(5), 966 – 991. https://doi.org/10.1037/edu0000685

Lang, F. R., Kamin, S., Rohr, M., Stünkel, C., & Williger, B. (2014). *Erfassung der fluiden kognitiven Leistungsfähigkeit über die Lebensspanne im Rahmen des Nationalen Bildungspanels: Abschlussbericht zu einer NEPS-Ergänzungsstudie* [Measuring the fluid cognitive ability across the lifespan in the National Educational Panel Study: Final report of a NEPS supplementary study]. *NEPS Working Paper*, Article 43. Retrieved from https://www.neps-data.de/Portals/0/Working%20Papers/WP_XLIII.pdf

Lenhard, W., & Schneider, W. (2006). *ELFE 1 – 6: Ein Leseverständnistest für Erst- bis Sechstklässler* [ELFE 1 – 6: A reading ability test for first to sixth graders]. Hogrefe.

Liu, J., Peng, P., Zhao, B., & Luo, L. (2022). Socioeconomic status and academic achievement in primary and secondary educa-

tion: A meta-analytic review. *Educational Psychology Review, 34*, 2867 – 2896. https://doi.org/10.1007/s10648-022-09689-y

Lindqvist, A., Björklund, F., & Bäckström, M. (2017). The perception of the poor: Capturing stereotype content with different measures. *Nordic Psychology, 69*(4), 231 – 247. https://doi.org/10.1080/19012276.2016.1270774

Lorenz, G. (2021). Subtle discrimination: Do stereotypes among teachers trigger bias in their expectations and widen ethnic achievement gaps? *Social Psychology of Education, 24*(2), 537 – 571. https://doi.org/10.1007/s11218-021-09615-0

Lorenz, C., Berendes, K., & Weinert, S. (2017). *Measuring receptive grammar in Kindergarten and elementary school children in the German National Educational Panel Study* (NEPS Survey Paper No. 24). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://doi.org/10.5157/NEPS:SP24:1.0

Lorenz, G., Gentrup, S., Kristen, C., Stanat, P., & Kogan, I. (2016). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen [Stereotypes among teachers? A study of systematic bias in teacher expectations]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 68*(1), 89 – 111. https://doi.org/10.1007/s11577-015-0352-3

Lorenz, G., Kogan, I., Gentrup, S., & Kristen, C. (2023). Nonnative accents among young students and teacher expectations for future student achievements. *Sociology of Education*, *97*(1), 76 – 96. https://doi.org/10.1177/00380407231202978

Lühe, J., Becker, M., Neumann, M., & Maaz, K. (2017). Geschlechtsspezifische Leistungsunterschiede in Abhängigkeit der sozialen Herkunft. Eine Untersuchung zur Interaktion zweier sozialer Kategorien [Gender-specific differences in achievement regarding social background: An examination of the interaction between two social factors]. *Zeitschrift für Erziehungswissenschaft, 20*, 499 – 519. https://doi.org/10.1007/s11618-016-0720-4

Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology, 72*, 791 – 809. https://doi.org/10.1037/0022-3514.72.4.791

Meissel, K., Meyer, F., Yao, E. S., & Rubie-Davies, C. M. (2017). Subjectivity of teacher judgments: Exploring student characteristics that influence teacher judgments of student ability. *Teaching and Teacher Education, 65*, 48 – 60. https://doi.org/10.1016/j.tate.2017.02.021

Mensah, F. K., & Kiernan, K. E. (2010). Gender differences in educational attainment: Influences of the family environment. *British Educational Research Journal, 36*(2), 239 – 260.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019. International results in mathematics and science.* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/international-results/

Mullis, I. V. S., von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021: International results in reading.* Retrieved from Boston College, TIMSS & PIRLS International Study Center. https://doi.org/10.6017/lse.tpisc.tr2103.kb5342

Muntoni, F., & Retelsdorf, J. (2018). Gender-specific teacher expectations in reading: The role of teachers' gender stereotypes. *Contemporary Educational Psychology, 54*, 212 – 220. https://doi.org/10.1016/j.cedpsych.2018.06.012

Nguyen, C. D., Carlin, J. B., & Lee, K. J. (2017). Model checking in multiple imputation: An overview and case study. *Emerging Themes in Epidemiology, 14*(8), 1 – 12. https://doi.org/10.1186/s12982-017-0062-6

Olczyk, M., Gentrup, S., Schneider, T., Volodina, A., Perinetti Casoni, V., Washbrook, E., Kwon, S. J., & Waldfogel, J. (2023). A cross-country perspective on teacher judgements and gender achievement gaps in primary education, *Social Science Review, 116*, Article 102938. https://doi.org/10.1016/j.ssresearch.2023.102938

Olczyk, M., Kwon, S. J., Lorenz, G., Perinetti Casoni, V., Schneider, T., Volodina, A., Waldfogel, J., & Washbrook, E. (2023). Teacher judgements, student social background, and student progress in primary school: A cross-country perspective. *Zeitschrift für Erziehungswissenschaft, 26*, 443 – 468. https://doi.org/10.1007/s11618-022-01119-7

Parker, P. D., Van Zanden, B., Marsh, H. W., Owen, K., Duineveld, J. J., & Noetel, M. (2020). The intersection of gender, social class, and cultural context: A meta-analysis. *Educational Psychology Review, 32*, 197 – 228. https://doi.org/10.1007/s10648-019-09493-1

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*(1), 239 – 253.

Petsko, C. D., & Bodenhausen, G. V. (2020). Multifarious person perception: How social perceivers manage the complexity of intersectional targets. *Social and Personality Psychology Compass, 14*(2), Article e12518. https://doi.org/10.1111/spc3.12518

Petsko, C. D., Rosette, A. S., & Bodenhausen, G. V. (2022). Through the looking glass: A lens-based account of intersectional stereotyping. *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes, 123*(4), 763 – 787. https://doi.org/10.1037/pspi0000382

Rauer, W., & Schuck, K. D. (2004). *FEESS 1 – 2: Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen* [FEESS 1 – 2: Questionnaire for emotional and social school experiences of first- and second-grade students]. Beltz Test.

Ready, D. D., & Chu, E. M. (2015). Sociodemographic inequality in early literacy development: The role of teacher perceptual accuracy. *Early Education and Development, 26*(7), 970 – 987. https://doi.org/10.1080/10409289.2015.1004516

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. https://doi.org/10.3102/0002831210374874

Rosén, M., Steinmann, I., & Wernersson, I. (2022). Gender differences in school achievement. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (pp. 1 – 48). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_46-1

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* Wiley.

Sachse, K. A., Jindra, C., Schumann, K., & Schipolowski, S. (2022). Soziale Disparitäten [Social disparitites]. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 151 – 180). Waxmann. https://doi.org/10.31244/9783830996064

Schnittjer, I., & Gerken, A.-L. (2018). *NEPS technical report for mathematics: Scaling results of Starting Cohort 2 for Grade 2* (NEPS Survey Paper No. 47). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://doi.org/10.5157/NEPS:SP47:1.0

Schuchardt, C., & Dunkake, I. (2014). Schichtspezifische Stereotype unter angehenden Lehrkräften [Social class stereotypes among teacher trainees]. *Zeitschrift für Soziologie der Erziehung und Sozialisation, 34*(1), 89 – 107.

Shepherd, M. A. (2011). Effects of ethnicity and gender on teachers' evaluation of students' spoken responses. *Urban Education, 46*(5), 1011 – 1028. https://doi.org/10.1177/0042085911400325

Sidanius, J., Hudson, S.-K. T. J., Davis, G., & Bergh, R. (2018). The theory of gendered prejudice: A social dominance and intersectionalist perspective. In A. Mintz, & L. Terris (Eds.), *The Oxford handbook of behavioral political science.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190634131.013.11

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*, 207 – 218.

StataCorp. (1985 – 2023). *Stata – statistics/data analysis.* Statistical software. StataCorp LLC.

StataCorp. (2023). *Stata multiple-imputation reference manual.* Stata Press. Retrieved from https://www.stata.com/manuals/mi.pdf

Steffens, M. C., & Jelenec, P. (2011). Separating implicit gender stereotypes regarding math and language: Implicit ability stereotypes are self-serving for boys and men, but not for girls and women. *Sex Roles, 64*, 324 – 335. https://doi.org/10.1007/s11199-010-9924-x

Stolier, R. M., & Freeman, J. B. (2016). The neuroscience of social vision. In J. R. Absher & J. Cloutier (Eds.), *Neuroimaging personality, social cognition, and character* (pp. 139 – 157). Academic Press. https://doi.org/10.1016/B978-0-12-800935-2.00007-5

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. https://doi.org/10.1037/a0027627

Tiedemann, J. (2002). Teacher's gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics, 50*, 49 – 62. https://doi.org/10.1023/A:1020518104346

Tobisch, A., & Dresel, M. (2020). Fleißig oder faul? Welche Einstellungen und Stereotype haben angehende Lehrkräfte gegenüber Schüler*innen aus unterschiedlichen sozialen Schichten? [Hardworking or lazy? What attitudes and stereotypes do prospective teachers have toward pupils from different social classes?]. In S. Glock & H. Kleen (Eds.), *Stereotype in der Schule.* Springer VS. https://doi.org/10.1007/978-3-658-27275-3_5

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*, 497 – 527. https://doi.org/10.3102/0002831209353594

Wang, S., Rubie-Davies, C. M., & Meissel, K. (2018). A systematic review of the teacher expectation literature over the past 30 years. *Educational Research and Evaluation, 24*(3 – 5), 124 – 179. https://doi.org/10.1080/13803611.2018.1548798

Weiß, R. H., & Osterland, J. (1997). *Grundintelligenztest Skala 1: CFT 1* [Culture fair intelligence test 1: CFT1]. Hogrefe.

Welling, J., & Zink, E. (2023). *NEPS Technical report for receptive grammar: Scaling results of Starting Cohort 2 in kindergarten and first grade* (NEPS Survey Paper No. 107). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://doi.org/10.5157/NEPS:SP107:1.0

Wenz, S. E., Olczyk, M., & Lorenz, G. (2016). *Measuring teachers' stereotypes in the NEPS* (NEPS Survey Paper No. 3). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://doi.org/10.2139/ssrn.2999297

White, I. R., Royston P., & Wood A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377 – 399.

Published online July 31, 2024

Open Data

The KuL data can be accessed through the Research Data Centre at the Institute for Educational Quality Improvement: https://www.iqb.hu-berlin.de/fdz/studies/KuL (DOI 10.5159/IQB_KuL_v1). The NEPS data can be accessed at the Leibniz Institute for Educational Trajectories (LIfBi): https://www.neps-data.de/Data-Center/Data-Access (DOI 10.5157/NEPS:SC2:10.0.0). The codes for the analysis of both datasets are available at the Open Science Framework (OSF): https://osf.io/2fsjk/

Funding

The authors acknowledge financial support from the Federal Ministry of Education and Research of Germany (project no. 01JC1117). The responsibility for the content of this publication lies with the authors.

Open access publication was enabled by Humboldt-Universität zu Berlin.

Dr. Sarah Gentrup
Department of Education Studies / IQB
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
sarah.gentrup@hu-berlin.de