Learning and Individual Differences 87 (2021) 101995

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Modifying biased teacher expectations in mathematics and German: A teacher intervention study

Markus P. Neuenschwander *, Camille Mayland, Edith Niederbacher, Ariana Garrote

University of Applied Sciences and Arts Northwestern Switzerland

ARTICLE INFO

Keywords:
Teacher expectations
Migration background
Socioeconomic status
Intervention study
Teacher training

ABSTRACT

Students' migration backgrounds and low socioeconomic status can bias teacher expectations of student achievement in mathematics and German. The main goal of this intervention study was to inform, raise awareness, and provide opportunities to implement behaviors to modify primary school teachers' biased achievement expectations. Before and after the implementation of the teacher training, data were collected using teacher, student, and parent questionnaires and student achievement tests in mathematics and German. Regression analyses using a sample of 860 students from Grades 4 to 6 from 75 classes showed that students' migration backgrounds and socioeconomic status biased teacher expectations for pretests in mathematics and German. After the intervention, expectations in mathematics were unbiased by students' migration background among teachers in the test group, in contrast to the control group. This study provides evidence for strategies to modify biased teacher expectations through teacher training.

1. Introduction

In the school context, the accuracy of teacher expectations is important (Tobisch & Dresel, 2017) because teacher expectations can predict students' academic achievement (Jussim, Eccles, & Madon, 1996; Niederbacher & Neuenschwander, 2020) and students' grades (Bonefeld & Dickhäuser, 2018) and can affect students' school careers (Ditton & Krüsken, 2009; Neuenschwander, Fräulin, Schumann, & Jüttler, 2018). Inaccurate low teacher expectations can influence student learning inappropriately and have a negative effect on students' academic achievement (Gentrup, Lorenz, Kristen, & Kogan, 2020). More importantly, students' socioeconomic status (SES) and migration backgrounds can bias teachers' expectations, which means teacher expectations tend to be inaccurately lower for students with low SES (Carigiet Reinhard, 2012; Lorenz, 2018) and migration backgrounds (Lorenz, Gentrup, Kristen, Stanat, & Kogan, 2016). In light of the negative effects that low teacher expectations can have, examining biased teacher expectations and investigating ways to modify them is important. Prior intervention studies have investigated the effects of increasing teachers' expectations (e.g., Timperley & Phillips, 2003) and the effects of behaviors associated with high-expectation teachers on student achievement and motivation (e.g., Rubie-Davies, 2015). However, little research exists on effective strategies to modify biased teacher expectations (de Boer, Timmermans, & van der Werf, 2018; Rubie-Davies & Rosenthal, 2016). The present study contributes to this research field by presenting an investigation of biased teacher expectations and evaluation of the effects of an intervention aimed at modifying biased teacher expectations in mathematics and German.

1.1. Biased teacher expectations and student social background

Teacher expectations are beliefs about future student achievement in a school subject (Ludwig, 2006). Teacher expectations are accurate if they correspond to students' actual achievement in a school subject (Jussim et al., 1996). In general, teachers base their expectations on students' achievement (Jussim et al., 1996; Südkamp, Kaiser, & Möller, 2012) and let students' effort influence them (Lorenz, 2018; Wang, Rubie-Davies, & Meissel, 2018). In addition, teachers perceive highly motivated students as competent and high achieving (Jussim et al., 1996).

Teachers not only base their expectations on students' achievement and effort; they can also be influenced by social stereotypes (Ganter, 1997). Social stereotypes are beliefs about social groups (e.g., migrants or citizens with low SES) that are largely and collectively internalized by

* Corresponding author at: University of Applied Sciences and Arts Northwestern Switzerland, Center for Learning and Socialization, Bahnhofstrasse 6, CH-5210 Windisch, Switzerland.
E-mail address: markus.neuenschwander@fhnw.ch (M.P. Neuenschwander).

https://doi.org/10.1016/j.lindif.2021.101995
Received 6 September 2020; Received in revised form 18 March 2021; Accepted 20 March 2021

society. When teachers activate social stereotypes, their perceptions of students correspond with their beliefs about the social group (Fiske & Neuberg, 1990; Reyna, 2008).

For instance, findings showed that students whose family language did not correspond with the instruction language (i.e., German) had lower achievement in German (Hippmann, Jambor-Fahlen, & Becker-Mrotzek, 2019). An active stereotype can lead teachers to expect lower achievement levels in the instruction language (i.e., German) for students with migration backgrounds than is accurate for those students (Appel, Weber, & Kronberger, 2015; Sander, Ohle, McElvany, Zander, & Hannover, 2018). Typically, this subject-specific (language) stereotype is generalized to other school subjects, such as mathematics (Appel et al., 2015). As a result, teachers underestimate the academic achievement level of students with migration backgrounds in various school subjects (Carigiet Reinhard, 2012; Gentrup et al., 2020).

However, teacher expectations vary by ethnic group (Lorenz et al., 2016). Migration-based teacher expectation can be confounded by students' SES (Makarova, 2008). Some ethnic groups tend to have low SES, whereas other ethnic groups tend to have high SES. Thus, effects of migration background can be interpreted better after controlling for SES. SES indicates an individual's or a family's ranking in a societal hierarchy based on access to a combination of valued commodities such as wealth, power, and social status (Mueller & Parcel, 1981). In addition, it comprises a set of family resources and a pattern of the children's socialization (Mueller & Parcel, 1981). Thus, students with low SES are often stereotyped as having low cognitive abilities and low achievement levels in mathematics and German (Reyna, 2000; Spencer & Castano, 2007), and teachers tend to underestimate their achievement levels in mathematics and German (Carigiet Reinhard, 2012; Lorenz, 2018).

1.2. Stability of teacher expectations

Kuklinski and Weinstein (2000) demonstrated that teacher expectations are stable over time. Fazio and Sherry (2020) showed that teachers can maintain stable expectations toward students, even if they receive information repeatedly on the inaccuracy of their expectations. Dweck (1999) suggested that the stability of ability beliefs depends on people's implicit theory of ability. Ability can be perceived as stable and not alterable (entity theory) or as malleable and developing (incremental theory). This means that teacher expectations are more likely modifiable if teachers perceive ability as malleable and developing. Blackwell, Trzesniewski, and Dweck (2007) reported in an intervention study that ability beliefs could be modified to become malleable. Thus, an intervention to modify teacher expectations should affect the teachers' beliefs regarding ability that underlie their expectations (de Boer et al., 2018).

1.3. Modification of biased teacher expectations

Prior research has shown that contact with a person of another social group can reduce the effects of activated stereotypes on expectations (e.g., Pettigrew & Tropp, 2008). Contact with members of another social group enhances individuals' (i.e., teachers') knowledge about the group and reduces their anxieties and increases their empathy concerning that social group. This helps reduce the effects of stereotyping. Thus, teacher expectations can become less biased when teachers have frequent contact with low-SES students and students with migration backgrounds (Glock, 2016). This finding is in line with research that showed teachers could modify their expectations by comparing low-SES and migrant students at their schools with a similar cohort at other schools (Perrella, 2017; Smith, 2007; Timperley & Phillips, 2003). Teachers learned that their low-SES and migrant students did not have lower achievement levels but performed at the same level as similar students enrolled in other schools. Reflecting on the achievement of low-SES students and of students with migration backgrounds in their classrooms helped teachers modify their preconceived, stereotyped expectations toward students and establish accurate expectations.

Intervention studies have presented additional detailed knowledge on how teacher expectations can be changed. Based on prior intervention studies on expectancy in educational contexts, Yeager and Walton (2011) analyzed which conditions in intervention studies were beneficial for the modification of expectations. They suggest that the modification of expectations can be effective and sometimes even long lasting when interventions strongly referred to the subjective experience of the target group (e.g., teachers) and when the new information was convincing.

Modification of teacher expectations had been investigated in only a few intervention studies, all of which were carried out in the United States, New Zealand, or China. In their review, de Boer et al. (2018) identified intervention studies that aimed at increasing low teacher expectations in general (e.g., Rubie-Davies, Peterson, Sibley, & Rosenthal, 2015). Other studies referred to specific groups of students such as students with low SES, minority students, low-achieving students, and students at risk for dropping out of school (e.g., Hui & Rubie-Davies, 2019; Timperley & Phillips, 2003; Weinstein et al., 1991). de Boer et al. (2018) identified three approaches to modifying teacher expectations through interventions that were often combined: (a) Teachers were instructed to apply behaviors associated with having high expectations. (b) Teachers were made aware of the effects of teacher expectations on students and that teacher expectations can be inaccurate and biased toward particular groups of students, such as students with low SES. (c) The underlying beliefs of teachers' biased achievement expectations toward student were addressed.

In line with the first and second approaches, Rubie-Davies et al. (2015) introduced in their study a teacher intervention centered around three core "high expectation teacher practices" to influence student achievement positively: (a) working in mixed-ability groups, (b) fostering a warm and supportive class climate, and (c) setting clear student goals focused on the mastery of skills. Teachers using these practices contributed to higher student achievement. Rubie-Davies et al. (2015) showed that teachers could learn those practices in workshops and apply the practices in their classes. The effects of this intervention study on student achievement were demonstrated in a randomized controlled trial. Students of teachers in the intervention group significantly improved their achievement in mathematics but not in reading.

Weinstein et al. (1991) conducted a pre-experimental study to modify teacher expectations for low achievers to prevent school failure. The intervention aimed at creating a positive school climate to improve student achievement and behavior at school. The program included changes in the school culture and in the curriculum, grouping, evaluation, motivation, student responsibility, and relationships in the classroom, with parents. Collaboration within school and with parents was an important factor that made the program effective. The intervention raised student motivation and student grades, but the improved achievement was not maintained.

Kerman (1979) developed the Teacher Expectation and Student Achievement program, which was evaluated in several studies. Hui and Rubie-Davies (2019) conducted one such study with a group of eight teachers in China who were made aware of differential treatment and its effects on students. They were instructed to implement behaviors in their classrooms, categorized according to three strands: challenging tasks, detailed feedback, and personal regard. They applied games and role-playing techniques to create awareness and exercised practices. The data were analyzed for the entire sample of students and for three subsamples of students with high, medium, and low teacher expectations. The findings showed that the teachers' behavioral change in their classrooms had a positive effect on the achievement gains of students. Additionally, the intervention affected self-conception among students for whom teachers had medium and low expectations but not among students for whom teachers had high expectations.

Perrella (2017) evaluated a U.S. Primary Talent Development Program on teacher perception of giftedness in culturally and linguistically

diverse students using semi-structured interviews. The program gave teachers strategies to identify gifted pupils in their school. Teachers used methods such as observation, parent or student surveys, alternative assessments, student portfolios, juried performances, standardized test scores, and teacher recommendations. Teachers reflected stereotypes and preconceived ideas of ability, and they learned that their preconceptions of students' abilities could limit students' career opportunities. As reported by Smith (2007), teachers changed their expectations of student achievement levels during the program.

In line with de Boer et al.'s (2018) second and third approaches, Timperley and Phillips (2003) implemented a program in schools in low-income communities to raise teacher expectations by giving information about and creating awareness of the students' ability to attain high achievement goals. Teachers were made aware that low-income students were able to attain a high level of reading achievement. In addition, researchers implemented activities that helped teachers monitor student achievement. Interviews with the teachers revealed that their expectations increased when student achievement increased.

In line with de Boer et al.'s (2018) third approach, Smith's (2007) pre-experimental study provided a similar finding. The results indicated that when teachers received information in the form of research papers, they did not change their beliefs. When teachers received data showing the actual achievement levels of their transient students was the same as the achievement levels of students in other schools, they changed their expectations. This suggests that teachers with reliable information on student achievement may change their expectation toward students.

The intervention studies discussed provide information about strategies for modifying teachers' expectations. All studies used a combination of strategies and methods to attain their goals. Some methods were implemented to raise teacher expectations in general or were focused on specific groups of students. Other methods aimed to improve student outcomes such as achievement, motivation, and self-concept or reduce the risk of dropping out of school. Several pre-experimental and experimental studies have addressed teacher expectations, but none aimed at modifying biased teacher expectations using an experimental design. Thus, increasing specific knowledge on how to modify teacher expectations biased by low SES and migration background is important.

1.4. Current study

In the current study, an intervention program that included three main strategies based on previous analyses and intervention studies was developed. First, based on the analyses of Yeager and Walton (2011) and Rubie-Davies (2015), teachers were informed of theories on teacher expectations and biased teacher expectation. Second, in line with Timperley and Phillips (2003), teachers were made aware of biased expectations toward students with low SES and students with migration backgrounds. Third, in accordance with prior studies (e.g., Weinstein et al., 1991), the information was related to their students and the program was implemented in their classrooms. All three strategies were applied in reference to students with low SES and migration backgrounds. The program was implemented using various methods such as role-playing, group discussions, theory presentations, reports by adults from low SES or migrant families, reflection about the students' abilities in their classes, cooperation with low-SES parents and migrant parents, and observation of teacher behaviors in the classroom, including feedback on this behavior.

The current study aimed to examine whether students' low SES and migration backgrounds biased teacher expectations. Prior studies have shown that teachers base their expectations not only on students' achievement but also on students' effort (Lorenz, 2018). Therefore, students' prior achievement and students' effort were controlled for in this study.

Two hypotheses were tested:

Hypothesis 1. Students' migration backgrounds predict subject-specific teacher expectations, after controlling for students' achievement and effort.

Hypothesis 2. Students' SES predicts subject-specific teacher expectations, after controlling for students' achievement and effort.

The effects of the intervention to modify biased teacher expectations in mathematics and German were evaluated. Two hypotheses focused on the intervention's effects on teacher expectations:

Hypothesis 3. The intervention would reduce teachers' biased subject-specific expectations of students with migration backgrounds.

Hypothesis 4. The intervention would reduce teachers' biased subject-specific expectations of students with low SES.

Though the concept of biased teacher expectations was introduced as a psychological process that affects various school subjects (Jussim et al., 1996), social stereotypes can be domain specific and interact with school subjects. Although students' migration backgrounds are strongly related to estimations of their language skills (Sander et al., 2018), students' SES is strongly related to estimations of cognitive abilities and competence in mathematics (Auwarter & Aruguete, 2008; Spencer & Castano, 2007). Therefore, all hypotheses were tested separately for mathematics and for German.

2. Method

2.1. Participants

The hypotheses were tested in a randomized controlled trial of a teacher expectation intervention that used a pretest–posttest design.

2.1.1. Teachers

Seventy-five teachers from 42 schools in six Swiss German-speaking cantons agreed to participate in the study with their classes. At the pretest (t1), 73 teachers filled out questionnaires on their students and 69 teachers filled out teacher questionnaires. Teachers in the intervention group ($n = 28$, 60.7% female) were 24 to 58 years old ($M = 35.61$, $SD = 10.51$). Most teachers (92.9%) were of Swiss nationality. Teachers in the control group ($n = 41$, 75.6% female) were 23 to 66 years old ($M = 39.95$, $SD = 14.52$) and most (87.8%) were of Swiss nationality. For six teachers, gender and nationality were missing. Seven teachers in the intervention group and two teachers in the control group left the study along with their classes because of personal reasons related to the teachers. For the posttest (t2), 22 teachers in the intervention group and 44 teachers in control group and their classes participated.

2.1.2. Students

From the 75 classes in Grades 4 to 6, 1480 students were asked to participate in the study and 1152 students agreed to participate. Students whose German skills were estimated by the teachers as too low to be able to participate in the assessment were excluded from the study ($n = 11$ students). Students in the intervention group ($n = 460$, 49.1% female) were 9 to 13 years old ($M = 11.10$, $SD = 0.93$). Around 47% of the students reported being of Swiss nationality ($n = 211$), 27.2% reported having dual citizenship (Swiss and other nationality, $n = 122$), and 26% reported being of foreign nationality ($n = 116$, missing $n = 11$). Students in the control group ($n = 692$, 51.2% female) were 8 to 13 years old ($M = 10.50$, $SD = 1.02$). Around 57% were of Swiss nationality ($n = 379$), 20.2% had dual citizenship ($n = 135$), and 23.1% were of foreign nationality ($n = 154$, missing $n = 24$). In line with Swiss population data, students of foreign nationality (23.4%) came from 56 different countries, most frequently Balkan countries (24.2%), Italy (9.6%), Germany (9.2%), and Turkey (7%).

Students with dual citizenship were excluded from the analyses for two reasons. First, categorization into one of two groups (i.e., with vs. without a migration background) did not adequately reflect students'

complex migration backgrounds. Second, students with dual citizenship differ significantly from both Swiss students and students of foreign nationality in terms of language competency (Dubowy et al., 2011; Wendt & Schwippert, 2017). Students with dual nationality ($n = 257$) or with missing values for nationality ($n = 35$) were removed (total $n = 292$). Thus, the student sample consisted of 860 students (327 in the intervention group and 533 in the control group).

2.1.3. Parents

Most of the 1152 students' parents filled out the parent questionnaire (435 in the intervention group and 663 in the control group). The first parental person of reference ($n = 1080$; missing = 18) was the mother (73.3%), the father (20.1%), or another adult (0.4%). The second parental person of reference ($n = 1060$; missing = 38) was the mother (19.4%), the father (65.9%), or another adult (4.2%). Data from 238 parents whose children had dual nationality were removed. A sample of 860 parents was analyzed.

2.2. Procedure

Swiss schools are regulated at the canton level and thus vary by canton in educational policy and curricula (Konsortium Überprüfung der Grundkompetenzen, 2016). From six German-speaking Swiss cantons, 756 randomly selected primary schools that were randomly assigned to intervention group or control group were asked to participate in the study. Forty-two school principals agreed to participate with their schools: 18 schools belonged to the intervention group and 24 schools to the control group. All teachers who taught mathematics or German in fourth, fifth, or sixth grade and who agreed to participate in the study were included. The acceptance rate varied between cantons and between the intervention and control group. Both groups were informed about the study's general goal (to gain knowledge on educational equality) to adhere to ethical guidelines, but the control group teachers were not informed about the intervention. In Switzerland, continuing education is compulsory for all teachers in most cantons. Thus, it can be assumed that the control group teachers were involved in professional programs at the time of the study. However, only the teachers in the intervention group received the special training of the intervention.

The study was conducted according to and in line with the guidelines of the research ethics board of the affiliated university. The guidelines require formal approval of proposed research if certain criteria (e.g., health studies) are fulfilled. The present study did not fulfill the criteria and thus did not require formal approval from the board. School principals, teachers, and parents were asked for their written informed consent. All participants voluntarily participated in the study. Teachers, parents, and students filled out the questionnaires at the beginning (t1, weeks 5–17) and the end (t2, weeks 28–39) of the 2016–2017 school year.

2.3. Description of the intervention

The intervention program was based on three evidence-based strategies implemented through four separate workshops and individual coaching by an expert. Teachers were trained in three groups that were about equal in size. The four workshops were standardized. Between the workshops (at intervals of approximately 1.5 months), the teachers implemented the intervention in their classrooms.

At the first workshop (8 h), teachers were informed about the topics of stereotyping, heterogeneity, and discrimination against students with migration backgrounds and low SES (Allemann-Ghionda, Auernheimer, Grabbe, & Krämer, 2006). Subsequently, to raise awareness about discrimination, they participated in role-playing exercises with roles of persons experiencing discrimination, and they reflected upon their experiences (Thiagarajan & van den Bergh, 2016).

At the second workshop (4 h), to apply the theory to their classrooms, the teachers analyzed their expectations of their students by filling out worksheets and reflecting on their students' achievement levels. To establish a basis for a positive relationship and raise their expectations of those students, teachers were asked to focus on the students' strengths and resources (Rubie-Davies, 2015). In addition, teachers learned typical practices of teachers with high expectations such as (a) establishing a positive classroom climate (rules and rituals), (b) adapting the feedback culture (focusing on students' strengths), (c) reacting to individual students' needs, (d) including more teacher-supportive teaching content (enrichment), and (e) including effort and emotion (Rubie-Davies, 2007). Teachers were asked to implement the first two practices (i.e., classroom climate and feedback culture) in their classrooms.

The third workshop (8 h) focused on enhancing successful cooperation with parents with low SES or migration backgrounds (Epstein et al., 2009), because studies have shown that teachers' perceptions of these parents influence their expectations toward students (Eccles & Harold, 1996; Niederbacher & Neuenschwander, 2020; Sheridan et al., 2012). To create awareness and implement the theory to practice, teachers discussed their experiences of cooperation with parents, formulated experienced challenges, and developed practical strategies for successfully managing such situations in the future (Weinstein et al., 1991).

At the last workshop (4 h), adults who grew up in low-SES families or with migration backgrounds who had attained higher educational levels than their parents had were invited to share their stories. They talked about school factors in their biographies that positively influenced their successful educational careers (e.g., teachers having high expectations for them) and engaged in discussions with the participating teachers. This close contact with adults coming from low-SES families or with migration backgrounds was intended to create awareness of biased teacher expectations (Glock, 2016). In this workshop, teachers learned that students with low SES or migration backgrounds can have successful educational careers and reflected on the general stereotype that low SES and migration backgrounds are associated with low student ability.

Research has shown that the strategies taught in group training exercises are better implemented in practice and become more sustainable if teachers receive personal coaching (Hui & Rubie-Davies, 2019; Joyce & Showers, 1981). Between the first and the last workshop, an expert coach helped teachers implement the course methods in their classrooms. Using the cognitive coaching approach (Costa & Garmston, 2002), the coach observed one or two school lessons given by the teacher that focused on a task chosen by the teacher. Afterward, the coach and the teacher discussed the behavior observed by the coach. Moreover, pairs of teachers were asked to visit on another's classrooms and give feedback on teacher behavior that might express biased achievement expectations. After the classroom visit, the teachers described their experiences in a project diary for further individual reflection. These methods are helpful not only for creating teacher awareness of biased expectations toward students with low SES and migration backgrounds, but also for supporting teachers in the implementation of the theory in their classrooms.

2.4. Measures

2.4.1. Subject-specific teacher expectations

To reduce teachers' time expenditure during data collection, teacher expectations were assessed using one item (e.g., Lorenz et al., 2016; Praetorius, Berner, Zeinz, Scheunpflug, & Dresel, 2013). During t1 and t2, teachers were asked to assign a rating of 1 (*poor*) to 6 (*outstanding*) for each student, indicating their level of expectation for that student's semester-end achievement in mathematics or German (Table 1).

2.4.2. Student achievement and effort

Students' achievements in mathematics and German were measured at t1 and t2 with an achievement test for each grade. The curriculum-

Table 1
Psychometric properties of all variables across the intervention and control groups.

Variable	Intervention group (n = 327)				Control group (n = 533)			
	n	M	SD	Missing	n	M	SD	Missing %
T-Expectations$_{Mt1}$	312	4.54	1.25	4.6%	521	4.52	1.25	2.3%
T-Expectations$_{Gt1}$	310	4.42	1.23	5.2%	521	4.41	1.21	2.3%
T-Expectations$_{Mt2}$	243	4.64	1.24	25.7%	487	4.58	1.23	8.6%
T-Expectations$_{Gt2}$	244	4.49	1.24	25.4%	487	4.51	1.20	8.6%
S-Achievement$_{Mt1}$	315	1.63	1.52	3.7%	524	0.99	1.53	1.7%
S-Achievement$_{Gt1}$	314	0.54	0.99	4.0%	522	0.40	1.04	2.1%
S-Achievement$_{Mt2}$	235	2.29	1.54	28.1%	466	1.68	1.62	12.6%
S-Achievement$_{Gt2}$	235	0.90	1.16	28.1%	465	0.76	1.04	12.8%
S-Effort$_{Mt1}$	324	4.95	0.83	0.9%	531	4.95	0.85	0.4%
S-Effort$_{Gt1}$	324	4.98	0.85	0.9%	531	5.01	0.80	0.4%
S-Migration background$_{t1}$	116	35.5 %		0.0%	154	28.9 %		0.0%
P-SES$_{t1}$	287	57.52	21.40	12.2%	474	59.26	19.56	11.1%

Note. T = teacher data; S = student data; P = parent data; $_{G}$ = German; $_{M}$ = mathematics; $_{t1}$ = pretest; $_{t2}$ = posttest; SES = highest family SES.

valid test items were open and closed. The mathematics achievement test for Grades 4 and 5 included tasks involving basic operations, applied calculus, logical thinking, and problem solving. The test for the sixth grade also included tasks involving fractions and decimals. The German achievement tests for all class levels focused on vocabulary, grammar, text comprehension, and reading. The validity of the Grades 5 and 6 tests in German and mathematics was reported in a previous study (Neuenschwander, Rottermann, Scheffler, & Rösselet, 2014). The fourth-grade tests were developed specifically for this study and comprised tasks from Moser, Buff, Angelone, and Hollenweger (2011) as well as tasks from a standardized test (Roick, Gölitz, & Hasselhorn, 2004). Split-half reliability indicated good values: mathematics$_{t1}$: r_{SH} = 0.89 to 0.90; mathematics$_{t2}$: r_{SH} = 0.90 to 0.94; German$_{t1}$: r_{SH} = 0.86 to 0.90; German$_{t2}$: r_{SH} = 0.84 to 0.95.

The items on the achievement tests were scored dichotomously (0 = *incorrect* and 1 = *correct*). Subsets of items from the fourth- and fifth-grade tests were presented in the fifth- and sixth-grade tests (anchor item design). Based on the item response theory (Yen & Fitzpatrick, 2006), Haberman's (2009) linking method was used to compute test values that were on the same metric for fourth to sixth grade. Weighted likelihood estimates were calculated (Warm, 1989).

Students' efforts in mathematics and German were assessed with four items each (e.g., "I do my best in class"), based on Rieger et al. (2017). Students rated each item on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The reliability was good (intervention group: α_{German} = 0.84, $\alpha_{mathematics}$ = 0.82; control group: α_{German} = 0.84, $\alpha_{mathematics}$ = 0.81).

2.4.3. Students' migration background

Students' migration background was assessed with the following questions: "Are you Swiss? Are you a foreigner? If you are a foreigner, from which country do you come?" Students could indicate if they were Swiss. If they were foreign or had additional nationalities, they could write their nationalities. The students' answers were scored dichotomously (0 = *Swiss nationality* and 1 = *foreign nationality*). The question did not focus on country of birth, because in Switzerland, country of birth does not define citizenship; information about country of birth would only be relevant for first-generation immigrant students and would not differentiate between second-generation immigrant students and Swiss students.

2.4.4. Parents' SES

Students' SES was assessed in the parents' questionnaire with a question about the current occupations of the first and second parental persons of reference, based on the International Standard Classification of Occupations 2008. Each occupation was coded with an International Socio-Economic Index of Occupational Status value (Ganzeboom & Treiman, 2010). For these analyses, the highest value of the parents' occupations was used.

2.5. Analytical strategy

The comparability of all t1 variables between the intervention group and the control group was tested using multivariate analyses of variances (MANOVA). A nonsignificant effect or a low effect size η^2 indicate that groups are considered equal. The dichotomous migration background variable was compared between groups using a χ^2 test; the measure of effect size was φ. In a subsequent step, correlations between all study variables for intervention and control groups were calculated to describe the data. The degrees of correlation were compared between groups using Fishers z test, and Cohens' q was used as the measure of effect size (Lenhard & Lenhard, 2016).

To test Hypotheses 1 to 4, regression analyses were calculated with the data for the whole sample. As a precondition to testing intervention effects, response biases between t1 and t2 were analyzed with t-tests and effect sizes with Cohen's d (Lenhard & Lenhard, 2016). Students who participated at both measurement times (longitudinal sample) were compared to students who participated only in t1 (cross-sectional sample). Significantly higher mean values were found for students who participated at both measurement times (longitudinal sample) for mathematics achievement, $t(1,110) = -3.02$, $p < .01$ (two-tailed), and for teacher expectations in mathematics and German, $t_{German}(1100) = -2.04$, $p < .05$ (two-tailed), $t_{mathematics}(1095) = -2.43$, $p < .05$ (two-tailed). The effects were small (mathematics achievement: d = 0.24; teacher expectations: d_{German} = 0.18, $d_{mathematics}$ = 0.22). No significant differences were found between these two samples in students' achievement in German, students' effort, migration background, and SES. Thus, there was not a response bias.

Percentages of missing data in the intervention group ranged from 0.9% to 25.7% and for the control group from 0.4% to 12.8% (Table 1). Data were assumed to be missing at random. Therefore, the missing values were imputed five times in five datasets using the NORM 2.03 program for calculating regression analyses (Schafer & Olsen, 1998). This program has the advantage that missing values of all variables are simultaneously estimated. Imputed missing data allow for better estimates in regression analyses than using listwise deletion (Graham, 2009).

Regression analyses to test Hypotheses 1 to 4 were calculated using Mplus 8.2 (Muthén & Muthén, 2018). Mplus 8.2 analyzed the five imputed data sets simultaneously by calculating a combined value over them. The data were structured hierarchically because teachers assessed all students in their classrooms. Thus, standard errors were controlled for the multilevel structure by including "type = complex" in the Mplus syntaxes (Muthén & Muthén, 2018). Students' prior achievement and effort in German and mathematics were included as control variables. To examine whether the effect of students' migration backgrounds and SES

on teachers' achievement expectations differed between the intervention group and the control group, each of these paths was set to zero. A significant χ^2 value indicated a difference of the paths between the groups (Byrne, 2012).

3. Results

As a first step, the similarity of the intervention group and control group was examined by comparing all the variables used at t1 between the intervention group and control group. A MANOVA with the following dependent variables with metric scales was conducted: achievement mathematics, achievement German, teacher expectations mathematics, teacher expectations German, effort mathematics, effort German, and SES. The MANOVA indicated a significant difference between the groups (Hotelling trace $F = 6.3$, $df = 7699$, $p < .001$, partial $\eta^2 = 0.059$). Univariate effects indicated a significant difference between groups for achievement mathematics; the other included variables did not significantly differ between groups. After including the canton variable as a second factor in the MANOVA, the intervention and control groups no longer differed (Hotelling trace $F = 1.5$, $df = 7, 694$, $p = .16$, partial $\eta^2 = 0.015$). The factor canton was significant (Hotelling trace $F = 4.3$, $df = 35, 694$, $p < .001$, partial $\eta^2 = 0.042$). In line with prior studies, the results showed significant differences in math achievement between cantons (Konsortium Überprüfung der Grundkompetenzen, 2016). The number of migrant students did not differ between the intervention and control groups, $\chi^2 = 2.7$, $df = 1$, $p = .10$, $\varphi = 0.06$). Thus, one can assume the intervention group and control group were comparable.

3.1. Bivariate correlations with study variables

Bivariate Pearson correlations between the variables were estimated using SPSS version 25 (Table 2). Correlations, which differed in their levels of significance between the control group and the intervention group, were analyzed with Fischer's z test for differences between the samples. The results showed that teacher expectations in mathematics and German at t1 and t2 in the intervention group and the control group were significantly related to students' migration backgrounds, parents' SES, students' achievement in mathematics and German (t1, t2), and students' self-reported effort in mathematics and German (t1). The correlations of students' migration backgrounds and their achievement in mathematics (t1, t2) differed significantly between the control group and the intervention group ($z_{t1} = 3.13$, $p < .01$; $z_{t2} = 2.58$, $p < .05$), but the effect size of these differences in correlations was small across both measurement times (migration background × achievement mathematics$_{t1}$: Cohen's $q = 0.22$; migration background × achievement mathematics$_{t2}$: Cohen's $q = 0.19$). The other correlations did not differ between the control group and the intervention group. This is additional evidence that the intervention group and the control group were comparable.

3.2. Biased teacher expectations

Hypotheses 1 and 2 assumed students' migration backgrounds and SES influence subject-specific teacher expectations after controlling for students' achievement and effort. To test Hypotheses 1 and 2, regression analyses with data from the control group and intervention group were calculated (Table 3). The data from the two groups were analyzed together to increase the power of the analyses and because the groups were considered to be equal regarding teacher expectations. The analysis tested whether students' migration backgrounds and parents' SES predicted teacher expectations in mathematics and German at t1. The effects of students' prior achievement and effort were controlled for. Students' SES and migration backgrounds significantly predicted teacher expectations in mathematics and in German (Hypotheses 1 and 2 were supported for both subjects). Thus, teacher expectations were biased by students' SES and migration backgrounds at the beginning of the school year.

Table 3
Effect of students' migration backgrounds and SES on teachers' achievement expectations in mathematics and German on the pretest ($_{t1}$).

Variable	Model 1a T-Expectations $_{Mt1}$ (n = 860)	Model 2a T-Expectations $_{Gt1}$ (n = 860)
	β (SE)	*β (SE)*
S-Migration background$_{t1}$	−0.06* (0.03)	−0.07* (0.04)
P-SES$_{t1}$	0.17*** (0.03)	0.17*** (0.04)
S-Achievement$_{Mt1}$	0.44*** (0.04)	
S-Achievement$_{Gt1}$		0.50*** (0.03)
S-Effort$_{Mt1}$	0.27*** (0.03)	
S-Effort$_{Gt1}$		0.15*** (0.03)
R^2	37%	42%

Note. T = teacher data; S = student data; P = parent data; $_G$ = German; $_M$ = mathematics; $_{t1}$ = first measurement time; SES = highest family SES. Migration background: 0 = Swiss nationality, 1 = Foreign nationality.
* $p < .05$, one-tailed.
*** $p < .001$, one-tailed.

Table 2
Bivariate correlations across the intervention group and the control group.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. T-Expectations$_{Mt1}$	1	0.66**	0.79**	0.60**	0.54**	0.48**	0.58**	0.49**	0.26**	0.20**	−0.16**	0.37**
2. T-Expectations$_{Gt1}$	0.81**	1	0.55**	0.80**	0.46**	0.61**	0.44**	0.62**	0.17**	0.27**	−0.26**	0.43**
3. T-Expectations$_{Mt2}$	0.81**	0.66**	1	0.68**	0.50**	0.47**	0.58**	0.47**	0.32**	0.20**	−0.14*	0.32**
4. T-Expectations$_{Gt2}$	0.69**	0.80**	0.77**	1	0.37**	0.61**	0.42**	0.62**	0.21**	0.36**	−0.24**	0.37**
5. S-Achievement$_{Mt1}$	0.52**	0.44**	0.49**	0.42**	1	0.57**	0.81**	0.56**	0.09	0.10	−0.23**	0.28**
6. S-Achievement$_{Gt1}$	0.55**	0.61**	0.51**	0.60**	0.63**	1	0.61**	0.81**	0.08	0.23**	−0.31**	0.32**
7. S-Achievement$_{Mt2}$	0.59**	0.53**	0.54**	0.51**	0.82**	0.68**	1	0.65**	0.26**	0.24**	−0.24**	0.25**
8. S-Achievement$_{Gt2}$	0.53**	0.61**	0.50**	0.61**	0.60**	0.84**	0.66**	1	0.11	0.28**	−0.35**	0.38**
9. S-Effort$_{Mt1}$	0.39**	0.28**	0.38**	0.30**	0.19**	0.17**	0.22**	0.15**	1	0.65**	−0.09	0.11
10. S-Effort$_{Gt1}$	0.25**	0.28**	0.23**	0.31**	0.13**	0.22**	0.16**	0.22**	0.73**	1	−0.14*	0.10
11. S-Migration background$_{t1}$	−0.19**	−−0.26**	−0.30**	−0.29**	−0.01	−0.18**	−0.06	−0.22**	−0.12**	−0.07	1	−0.38**
12. P-SES$_{t1}$	0.26**	0.30**	0.28**	0.33**	0.20**	0.27**	0.23**	0.27**	0.05	0.02	−0.26**	1

Note. Above the diagonal: intervention group: ($227 < n < 324$); below the diagonal: control group ($225 < n < 521$). T = teacher data; S = student data; P = parent data; $_G$ = German; $_M$ = mathematics; $_{t1}$ = pretest; $_{t2}$ = posttest; SES = highest family SES. Migration background: 0 = Swiss nationality, 1 = foreign nationality.
* $p < .05$, two-tailed.
** $p < .01$, two-tailed.

3.3. Intervention effects on biased teacher expectations

To test Hypotheses 3 and 4, multigroup regression analyses were run (Table 4). The results in mathematics showed a significant effect of students' migration backgrounds on teacher expectations at t2 in the control group after controlling for teacher expectations at t1, student achievement, and student effort. If the path between students' migration backgrounds and teacher expectations in mathematics at t2 was constrained to be equal between the groups, the χ^2 value was significant ($\chi^2[1] = 10.8$, $p < .001$). Thus, the effect of students' migration backgrounds on teachers' expectations in mathematics differed significantly between the groups (Hypothesis 3 was supported for mathematics). If the path from parents' SES to teachers' expectations in mathematics at t2 was constrained to be equal between the groups, the χ^2 value was not significant ($\chi^2[1] = 0.09$, $p = .76$, Hypothesis 4 was rejected for mathematics).

If the path between teachers' expectations in German and students' migration backgrounds was constrained to be equal between the groups, the χ^2 value was not significant ($\chi^2[1] = 0.15$, $p = .70$, Hypothesis 3 was rejected for German). Further, the parents' SES predicted teachers' expectations in German at t2 in the control group, but not in the intervention group. After setting the path as equal between the groups, the χ^2 value was not significant ($\chi^2[1] = 0.24$, $p = .62$). Thus, the effect did not differ significantly between the groups (Hypothesis 4 was rejected for German).

4. Discussion

This study aimed to examine biased teacher expectations in German and mathematics and to modify biased teacher expectations by implementing an intervention. As hypothesized, the results showed that teachers underestimated the achievement of students with migration backgrounds and low SES at the pretest, which is in line with prior research (Carigiet Reinhard, 2012; Lorenz, 2018). In addition, the results supported the assumption that biased teacher expectations of students can be modified. The intervention positively influenced teacher expectations in mathematics, so they were no longer biased by students' migration backgrounds. Although the intervention effect was only found for migration background (not for SES) and only in mathematics (not in German), the fact that it was found in a control group study using quantitative methods and controlling for SES, prior expectations, and student achievement and effort enhances the findings' relevance.

This finding provides evidence for the effect of a combination of three strategies: informing teachers about theory on teacher expectations and about biased teacher expectations, creating teacher awareness of biased expectations toward students with low SES and students with migration backgrounds, and applying theory and information to teaching the students in the classrooms. However, whether some strategies were more effective than others cannot be determined. Although providing information could create an understanding of the concept of expectations, creating awareness could serve as a motivational factor prompting change. Applying the intervention's lessons to teaching their students makes the intervention concrete and convincing. Yeager and Walton (2011) and Rubie-Davies (2015) indicated that helping teachers apply the theory to their individual classrooms and students is a decisive factor in reducing bias and improving outcomes. Thus, the combination of providing information, creating awareness, and implementing changes in the classroom seems to be crucial (Yeager & Walton, 2011).

Contrary to expectations, the intervention did not modify teacher expectations biased by student SES. One possible explanation could be that the intervention workshops focused more on the topic of migration than they did on low SES. Workshops with greater focus on how low SES can bias teacher expectations might help modify teachers' low expectations (Timperley & Phillips, 2003). Another possible explanation is that the SES bias is more entrenched. In line with Lorenz et al. (2016), the presented findings show that SES is a stronger predictor of biased teacher expectations than migration background. This could make teacher expectations that are biased by student SES more difficult to modify.

By the end of the school year, migration background no longer biased teacher achievement expectations in German in the control group. However, the expectations in mathematics remained biased by migration background. Frequent contact with persons of a stereotyped group can change biased expectations in a domain, so expectations become accurate (Glock, 2016; Pettigrew & Tropp, 2008). Teacher expectations in German may rely on observations of students in class and at recess, while teacher expectations in mathematics are influenced only by observations of the student specifically while engaged in mathematics. Therefore, teachers might have formed more accurate evaluations of students' German skills and established more accurate achievement expectations in German (Smith, 2007), but not in mathematics. Participating in the study might have reinforced this process. Teachers in the control group were informed about the study's general goals. Though only general information was given, teachers could have been influenced to observe migrant students' German skills more carefully (Hawthorn effect; Coombs & Smith, 2003). Thus, the stereotype actively continued to bias teachers' expectations in mathematics, explaining the unchanged biased teacher expectations in mathematics at the posttest in contrast to the unbiased teacher expectations in German at the posttest.

Moreover, in the control group, SES no longer biased teacher

Table 4
Effect of students' migration backgrounds and SES on teachers' achievement expectations in mathematics and German on the Posttest ($_{t2}$) after an intervention (Multigroup Regression Analyses).

Variable	Model 1a intervention T-Expectations $_{Mt2}$ ($n = 327$)	Model 1b control T-Expectations $_{Mt2}$ ($n = 533$)	Model 2a intervention T-Expectations $_{Gt2}$ ($n = 327$)	Model 2b control T-Expectations $_{Gt2}$ ($n = 533$)
	β (*SE*)	β (*SE*)	β (*SE*)	β (*SE*)
S-Migration background$_{t1}$	0.01 (0.04)	−0.12**(0.03)	0.02 (0.03)	−0.03 (0.03)
P-SES$_{t1}$	0.04 (0.04)	0.04 (0.03)	0.07 (0.06)	0.06* (0.03)
T-Expectations$_{Mt1}$	0.68*** (0.04)	0.69*** (0.04)		
T-Expectations$_{Gt1}$			0.64*** (0.05)	0.64***(0.04)
S-Achievement$_{Mt2}$	0.15* (0.05)	0.12*** (0.04)		
S-Achievement$_{Gt2}$			0.17**(0.06)	0.18*** (0.03)
S-Effort$_{Mt1}$	0.07*(0.07)	0.06** (0.03)		
S-Effort$_{Gt1}$			0.11**(0.04)	0.09***(0.03)
R^2	67%	69%	67%	68%

Note. T = teacher data; S = student data; P = parent data; $_G$ = German; $_M$ = mathematics; $_{t1}$ = pretest; $_{t2}$ = posttest; SES = highest family SES. Migration background: 0 = Swiss nationality, 1 = Foreign nationality.
* $p < .05$, one-tailed.
** $p < .01$.
*** $p < .001$.

expectations in mathematics at the posttest, but SES continued to bias teacher expectations in German at the posttest. The modification of teacher expectations in mathematics may have resulted from teacher observations of their low-SES students in mathematics and adapted their expectations accordingly. This assumption needs further investigation in future studies.

The study has several limitations. First, the intervention included three strategies. The collected data did not provide information about the individual strategies. The findings indicate only that the combination of the strategies was effective. Second, the schools and not the teachers were randomly assigned to either the intervention or the control group (Zhu, Jacob, Bloom, & Xu, 2012). Thus, when teachers agreed to participate with their classes, they already knew the study condition. A preparatory study showed that teachers were not willing to participate in a study without knowing whether they would have to participate in an intervention. In addition, the analyses indicated that the intervention group and the control group were similar. Thus, one can assume the data collection strategy did not affect the results. Third, the same achievement tests were used for the pretest and posttest, which could have resulted in practice effects. However, the probability that practice effects on the achievement test at the posttest affected the intervention effects is small, and they would affect both the intervention group and the control group. In addition, the analyses included achievement data from only one measurement point at a time (pretest or posttest). More specifically, the analyses of the intervention effects referred only to the achievement test at the posttest. Fourth, the study design did not allow examination of longitudinal effects of modified teacher expectations on student achievement. Additional studies should investigate the effects of interventions on biased teacher expectations and on student achievement. Fifth, the long-term effect of the intervention is unknown. Whether the teachers maintained their unbiased expectations toward migrant students in new classrooms would be interesting to know. Previous studies indicated that teachers may remain aware of the effects of their behavior after the implementation of an intervention (Hui & Rubie-Davies, 2019; Yeager & Walton, 2011). Sixth, this study focused on low-bias teacher expectations. Some teachers in the sample may have had inaccurately high expectations (Gentrup et al., 2020). The analytical strategy using continuous variables did not allow the distinction between inaccurately high and low teacher expectations. For example, finding out whether inaccurately high teacher expectations can be modified would be interesting. Finally, this study was conducted with teachers from different schools who taught in Grades 4–6. The intervention might have been more effective if all teachers of the participating schools were involved in a school-wide intervention (Neuenschwander & Niederbacher, 2019; Weinstein et al., 1991). Including all teachers of a school would probably promote a more intensive exchange among teachers and thus might increase the intervention effect.

5. Conclusion

In line with de Boer et al. (2018), a combination of multiple strategies can help modify biased teacher expectations. Providing information about the effects of (biased) teacher expectations in combination with creating awareness and being able to implement theories and behaviors in concrete classroom situations have proven effective (Yeager & Walton, 2011). Knowledge of how to modify biased teacher expectations can be applied in teacher education and in further teacher education. In addition, intervention programs should address the specific risks of students with low SES and with migration backgrounds. Finally, when modifying teacher expectations, the effects of their interactions with subject specificities need to be considered (Jussim et al., 1996). In line with prior intervention studies on teacher expectations (e.g., Rubie-Davies et al., 2015), the effects of the intervention differed between school subjects. Future research should continue to investigate the process of how teacher expectations in various subjects are determined and how they can be modified effectively to become accurate. Researchers and practitioners in the field need more knowledge of how biased teacher expectations can be positively influenced to improve the learning conditions of at-risk students.

Acknowledgements

The data collection and data analysis were supported by the Swiss Mercator Foundation (Grant Number 2014-0527), the State Secretariat for Migration, and the Swiss cantons Zürich and Aargau.

References

Allemann-Ghionda, C., Auernheimer, G., Grabbe, H., & Krämer, A. (2006). Beobachtung und Beurteilung in soziokulturell und sprachlich heterogenen Klassen [monitoring and assessment in socio-culturally and linguistically heterogeneous classes]. *Zeitschrift für Pädagogik, Beiheft, 51*, 250–266. https://doi.org/10.1007/978-3-531-18846-1_2.

Appel, M., Weber, S., & Kronberger, N. (2015). The influence of stereotype threat on immigrants: Review and meta-analysis. *Front. Psychol., 6*(900). https://doi.org/10.3389/fpsyg.2015.00900.

Auwarter, A. E., & Aruguete, M. S. (2008). Effects of student gender and socioeconomic status on teacher perceptions. *J. Educ. Res., 101*(4), 242–246. https://doi.org/10.3200/JOER.101.4.243-246.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Dev., 78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x.

de Boer, H., Timmermans, A. C., & van der Werf, M. P. C. (2018). The effects of teacher expectation interventions on teachers' expectations and student achievement: Narrative review and meta-analysis. *Educ. Res. Eval., 24*(3–5), 180–200. https://doi.org/10.1080/13803611.2018.1550834.

Bonefeld, M., & Dickhäuser, O. (2018). (Biased) grading of students' performance: Students' names, performance level, and implicit attitudes. *Front. Psychol., 9*, 481. https://doi.org/10.3389/fpsyg.2018.00481.

Byrne, B. M. (2012). *Multivariate Applications Series: Structural Equation Modeling with Mplus*. Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780203807644.

Carigiet Reinhard, T. (2012). *Schulleistungen und Heterogenität [Academic achievement and heterogeneity]*. Haupt.

Coombs, S. J., & Smith, I. D. (2003). The Hawthorne effect: Is it a help or a hindrance in social science research? *Change Transformations in Education, 6*, 97–111.

Costa, A. L., & Garmston, R. J. (2002). *Cognitive coaching: A foundation for renaissance schools*. Christopher-Gordon Publishers.

Ditton, H., & Krüsken, J. (2009). Bildungslaufbahnen im differenzierten Schulsystem—Entwicklungsverläufe von Laufbahnempfehlungen und Bildungsaspirationen in der Grundschulzeit [Educational careers in a tracked school system—Development of teacher recommendations and educational aspirations over the elementary school years]. *Z. Erzieh., 12*(1), 74–102. https://doi.org/10.1007/978-3-531-92216-4_4.

Dubowy, M., Duzy, D., Pröscholdt, M. V., Schneider, W., Souvignier, E., & Gold, A. (2011). Was macht den «Migrationshintergrund» bei Vorschulkindern aus? [what makes up the "migration background" of preschool children?]. *Schweizerische Zeitschrift für Bildungswissenschaften, 33*(3), 355–376.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Psychology Press. https://doi.org/10.4324/9781315783048.

Eccles, J. S., & Harold, R. D. (1996). Family involvement in children's and adolescents' schooling. In A. Booth, & J. F. Dunn (Eds.), *Family-school links: How do they affect educational outcomes?* (pp. 3–34). Lawrence Erlbaum.

Epstein, J. L., Sanders, M., Sheldon, S., Simon, B., Salinas, K., Jansorn, N., … Williams, K. J. (2009). *School, family, and community partnerships: Your handbook for action* (Corwin).

Fazio, L. K., & Sherry, C. L. (2020). The effect of repetition on truth judgments across development. *Psychol. Sci., 31*, 1150–1160. https://doi.org/10.1177/0956797620939534.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–74). Academic Press. https://doi.org/10.1016/S0065-2601(08)60317-2.

Ganter, S. (1997). *Stereotype und Vorurteile [stereotypes and prejudices]*. Mannheimer Zentrum für Europäische Sozialforschung.

Ganzeboom, H. B. G., & Treiman, D. J. (2010). *International stratification and mobility file: Conversion tools*. Department of Social Research Methodology.

Gentrup, S., Lorenz, G., Kristen, C., & Kogan, I. (2020). Self-fulfilling prophecies in the classroom: Teacher expectations, teacher feedback and student achievement. *Learn. Instr., 66*, 101–296. https://doi.org/10.1016/j.learninstruc.2019.101296.

Glock, S. (2016). Does ethnicity matter? The impact of stereotypical expectations on in-service teachers' judgments of student. *Soc. Psychol. Educ., 19*, 493–509. https://doi.org/10.1007/s11218-016-9349-7.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annu. Rev. Psychol., 60*, 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530.

Haberman, S. J. (2009). *Linking parameter estimates derived from an item response model through separate calibrations*. Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2009.tb02197.x.

Hippmann, K., Jambor-Fahlen, S., & Becker-Mrotzek, M. (2019). Der Einfluss familiärer Hintergrundvariablen auf die Leseleistung von Grundschulkindern im Anfangsunterricht [the impact of family background variables on the reading skills of children in elementary schools]. *Z. Erzieh., 22*(1), 51–73. https://doi.org/10.1007/s11618-018-0861-8.

Hui, D., & Rubie-Davies, C. M. (2019). Teacher expectation intervention: Is it effective for all students? *Learn. Individ. Differ., 74*, 101–751. https://doi.org/10.1016/j.lindif.2019.06.005.

Joyce, B., & Showers, B. (1981). *Transfer of training: The contribution of "coaching". 162 (2)* pp. 163–172). Boston University Journal of Education. https://doi.org/10.1177/002205748116300208.

Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Adv. Exp. Soc. Psychol., 28*, 281–388. https://doi.org/10.1016/S0065-2601(08)60240-3.

Kerman, S. (1979). *"Why did you call on me? I didn't have my hand up!": Teacher expectations and student achievement. 60* pp. 716–718). Phi Delta Kappan.

Konsortium Überprüfung der Grundkompetenzen. (2016). Überprüfung der Grundkompetenzen: Mathematik 11. In *Schuljahr [evaluation of basic competences: Mathematics 11th grade]*. EDK und SRED.

Kuklinski, M. R., & Weinstein, R. S. (2000). Classroom and grade level differences in the stability of teacher expectations and perceived differential teacher treatment. *Learn. Environ. Res., 3*, 1–34. https://doi.org/10.1023/A:1009904718353.

Lenhard, W., & Lenhard, A. (2016). Calculation of effect sizes. *Psychometrica.*. https://doi.org/10.13140/RG.2.1.3478.4245.

Lorenz, G. (2018). *Selbsterfüllende Prophezeiungen in der Schule [Self-fulfilling prophecies at school]*. Springer VS. https://doi.org/10.1007/978-3-658-19881-7.

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.

Ludwig, P. H. (2006). Erwartungen [Expectations]. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (pp. 144–150) (Psychologie Verlags Union).

Makarova, E. (2008). *Akkulturation und kulturelle Identität [acculturation and cultural identity]*. Haupt.

Moser, U., Buff, A., Angelone, D., & Hollenweger, J. (2011). *Nach sechs Jahren Primarschule [after six years of primary school]* (Bildungsdirektion Kanton Zürich).

Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Dev., 52*(1), 13–30. https://doi.org/10.2307/1129211.

Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide*.

Neuenschwander, M. P., Fräulin, J., Schumann, S., & Jüttler, A. (2018). Gymnasium oder Berufsausbildung? [gymnasium or vocational education and training?]. *Zeitschrift für Soziologie der Erziehung und Sozialisation, 38*(4), 364–382. https://doi.org/10.3262/ZSE1804361.

Neuenschwander, M. P., & Niederbacher, E. (2019). Förderliche und gerechte Beurteilung—der SCALA-Ansatz [Favorable and fair assessment—The SCALA approach]. *Schweizerische Zeitschrift für Heilpädagogik, 25*(7–8), 50–55.

Neuenschwander, M. P., Rottermann, B., Scheffler, L., & Rösselet, S. (2014). *Wirkungen der* Selektion *WiSel. Dokumentation der Leistungstests Mathematik und Deutsch der Welle 2* 2012/2013. In *(Forschungsbericht) [effects of selection—The project WiSel: Documentation of the achievement tests in mathematics and German, wave 2, 2012/2013 (research report)]*. Pädagogische Hochschule Nordwestschweiz.

Niederbacher, E., & Neuenschwander, M. P. (2020). Herkunftsbedingte Leistungsdisparitäten [Disparities in academic achievement due to social background]. *Z. Erzieh., 23*(4), 739–768. https://doi.org/10.1007/s11618-020-00955-9.

Perrella, C. J. (2017). *Exploring the influence of an early talent development program on teacher perceptions of giftedness in culturally and linguistically diverse students*. doctoral dissertation. Northeastern University https://repository.library.northeastern.edu/files/neu:cj82q089m/fulltext.pdf.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *Eur. J. Soc. Psychol., 38*, 922–934. https://doi.org/10.1002/ejsp.504.

Praetorius, A.-K., Berner, V.-D., Zeinz, H., Scheunpflug, A., & Dresel, M. (2013). Judgment confidence and judgment accuracy of teachers in judging self-concepts of students. *J. Educ. Res., 106*(1), 64–76. https://doi.org/10.1080/00220671.2012.667010.

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educ. Psychol. Rev., 12*(1), 85–110. https://doi.org/10.1023/A:1009037101170.

Reyna, C. (2008). Ian is intelligent but Leshaun is lazy: Antecedents and consequences of attributional stereotypes in the classroom. *Eur. J. Psychol. Educ., 23*, 439–458. https://doi.org/10.1007/BF03172752.

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). *Social cognitive constructs are just as stable as the big five between grades 5 and 8. 3(3)* pp. 1–9). American Educational Research Association Open. https://doi.org/10.1177/2332858417717691.

Roick, T., Gölitz, D., & Hasselhorn, M. (2004). DEMAT 3+. In *Deutscher Mathematiktest für dritte Klassen [DEMAT 3+. Mathematics test for third grades/classes. German version]*. Beltz Test GmbH.

Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high and low expectation teachers. *Br. J. Educ. Psychol., 77*, 289–306. https://doi.org/10.1348/000709906X101601.

Rubie-Davies, C. M. (2015). Becoming a high expectation teacher: Raising the bar. *Routledge.*. https://doi.org/10.4324/9781315761251.

Rubie-Davies, C. M., Peterson, E. R., Sibley, C. G., & Rosenthal, R. (2015). A teacher expectation intervention: Modelling the practices of high expectation teachers. *Contemp. Educ. Psychol., 40*, 72–85. https://doi.org/10.1016/j.cedpsych.2014.03.003.

Rubie-Davies, C. M., & Rosenthal, R. (2016). Intervening in teachers' expectations: A random effects meta-analytic approach to examining the effectiveness of an intervention. *Learn. Individ. Differ., 50*, 83–92. https://doi.org/10.1016/j.lindif.2016.07.014.

Sander, A., Ohle, A., McElvany, N., Zander, L., & Hannover, B. (2018). Stereotypenbedrohung als Ursache für geringeren Wortschatzzuwachs bei Grundschulkindern mit Migrationshintergrund [stereotype priming as a cause of lower vocabulary growth in primary school students with immigrant background]. *Z. Erzieh., 21*, 177–197. https://doi.org/10.1007/s11618-017-0763-1.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivar. Behav. Res., 33*(4), 545–571. https://doi.org/10.1207/s15327906mbr3304_5.

Sheridan, S. M., Bovaird, J. A., Glover, T. A., Garbacz, S. A., Witte, A., & Kwon, K. (2012). A randomized trial examining the effects of conjoint behavioral consultation and the mediating role of the parent–teacher relationship. *Sch. Psychol. Rev., 41*(1), 23–46. https://doi.org/10.1080/02796015.2012.12087374.

Smith, D. R. (2007). Changing teachers' beliefs about students and learning. In D. W. Lick, & C. U. Murphy (Eds.), *The whole-faculty study groups fieldbook: Lessons learned and best practices from classrooms, districts and schools* (pp. 196–201). Corwin Press.

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! *Soc. Justice Res, 20*, 418–432. https://doi.org/10.1007/s11211-007-0047-7.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *J. Educ. Psychol., 104*(3), 743–762. https://doi.org/10.1037/a0027627.

Thiagarajan, S., & Van den Bergh, S. (2016). *Interaktive Trainingsmethoden [interactive training methods]*. Wochenschau Verlag.

Timperley, H. S., & Phillips, G. (2003). Changing and sustaining teachers' expectations through professional development in literacy. *Teach. Teach. Educ., 19*, 627–641. https://doi.org/10.1016/S0742-051X(03)00058-1.

Tobisch, A., & Dresel, M. (2017). Negatively or positively biased? Dependencies of teachers' judgments and expectations based on students' ethnic and social backgrounds. *Social Psychology of Education and Training, 20*, 731–752. https://doi.org/10.1007/s11218-017-9392-z.

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

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450. https://doi.org/10.1007/BF02294627.

Weinstein, R. S., Soulé, C. R., Collins, F., Cone, J., Mehlhorn, M., & Simontacchi, K. (1991). Expectations and high school change: Teacher–researcher collaboration to prevent school failure. *Am. J. Community Psychol., 19*, 333–363. https://doi.org/10.1007/BF00938027.

Wendt, H., & Schwippert, K. (2017). Lesekompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund [students reading competences with and without migration background]. In A. Hussmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, … R. Valtin (Eds.), *IGLU 2016: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 219–234). Waxmann.

Yeager, D. S., & Walton, G. M. (2011). Social–psychological interventions in education: They're not magic. *Rev. Educ. Res., 81*(2), 267–301. https://doi.org/10.3102/0034654311405999.

Yen, W. M., & Fitzpatrick, A. R. (2006). Item response theory. In R. L. Brennan (Ed.), *Educational measurement*. Praeger.

Zhu, P., Jacob, R., Bloom, H., & Xu, Z. (2012). Designing and analyzing studies that randomize schools to estimate intervention effects on student academic outcomes without classroom-level information. *Educ. Eval. Policy Anal., 34*, 45–68. https://doi.org/10.3102/0162373711423786.