Negatively or positively biased? Dependencies of teachers' judgments and expectations based on students' ethnic and social backgrounds

Anita Tobisch, Markus Dresel

Negatively or positively biased? Dependencies of teachers' judgments and expectations based on students' ethnic and social backgrounds

Anita Tobisch[1] · Markus Dresel[1]

Abstract Based on theories of social information processing and judgment formation, we investigated whether teachers' achievement expectations, achievement aspirations and judgments of achievement-relevant characteristics depend on students' ethnic and/or social backgrounds. Furthermore, we addressed whether judgments for minority students are negatively biased or judgments for majority students are positively biased. To answer these questions, we conducted an online-study with 237 primary school teachers in Germany. We employed case vignettes and experimentally varied students' ethnic and social backgrounds by means of assigning specific first names. Teachers were asked to rate specific achievement expectations and achievement aspirations (grades for main subjects) for each student as well as provide judgments of achievement-relevant characteristics (general abilities, willingness to put in effort, qualification for a higher secondary school). Results from multi- and univariate analyses of variance with subsequent contrast analyses revealed significant differences in teachers' judgments for all considered characteristics dependent on students' ethnic and social backgrounds. Results suggested that teachers' achievement expectations and achievement aspirations are quite accurate for students with an immigrant background, but that teachers overestimate students without an immigration background and with high socioeconomic status. Findings are discussed with regard to automated and controlled information processes.

Keywords Achievement expectations · Teacher judgment · Socioeconomic status · Ethnic background

Anita Tobisch
anita.tobisch@phil.uni–augsburg.de

[1] Department of Psychology, University of Augsburg, Universitaetsstrasse 10, 86159 Augsburg, Germany

1 Introduction

Across several countries students' academic achievements are still associated with their ethnic and social backgrounds (e.g., Aud et al. 2010; Bradbury et al. 2015; Martin et al. 2012; Mullis et al. 2012; OECD 2012). The proportion of school dropouts is notably higher within ethnic minorities and students from low socioeconomic backgrounds (American Psychological Association 2012; Aud et al. 2010; Fuhr 2012; Hauser et al. 2000). Furthermore, such students have to repeat a school year more often than majority students from high socioeconomic backgrounds (Krohne et al. 2004). Both primary school and secondary school students from ethnic minorities and from low-income households achieve lower performances on international achievement tests (OECD 2016; Schwippert et al. 2012; Wendt et al. 2012). Although there is wide empirical evidence for disparities in academic performance dependent on ethnic and social backgrounds, there is still little known about possible reasons for such disparities in academic performance and the specific influences of students' ethnic and social backgrounds. In this context many factors are discussed, especially students' actual abilities and verbal skills, family conditions and support, the composition of class, teachers' backgrounds, and teachers' behaviors (e.g., American Psychological Association 2012; Dee 2005; Esser 2006; Kluczniok et al. 2013; Kristen 2002).

In order to plan and implement effective learning time and support each student's individual abilities, it is necessary that teachers accurately judge a student's performance and achievement-relevant dispositions. This essential task of the teaching profession is important to develop realistic expectations about students' future performances and behaviors. The reliability of school grades (Brimi 2011; Ingenkamp 1989; Praetorius et al. 2013) as well as teachers' judgment accuracy (McElvany et al. 2009; Südkamp et al. 2012) are frequently criticized as being low. Teachers' achievement expectations and judgments of achievement-relevant dispositions are, like other social judgments, vulnerable to cognitive biases. Especially if these biases are systematically linked with social categories, there might be the danger of disadvantaging some social groups, like ethnic or social minorities (e.g., Glock et al. 2013; Sprietsma 2013). These systematic cognitive biases may occur if fast decisions, that favor the use of heuristics, are needed. However, decisions and judgments are not only caused by automatic processes, but can also be more or less controlled.

Two-process-models of social judgments, like the continuum-model of impression formation (Fiske et al. 1999; Fiske and Neuberg 1990), explain how teachers form judgments and expectations about students either automatically (based on membership to social categories) or controlled (based on individual characteristics). Under some conditions (such as limited time or cognitive capacities) social information processes tend to be more automatic. In contrast, high motivation for information processing or personal relevance to the decision can lead to a more controlled process. As people prefer consistent information (Festinger 1957), their perception and interpretation of social categories is colored by their previous experiences and knowledge about social groups. Furthermore, social perception is

influenced by characteristics associated with these groups, which are manifested as mental representations in their memory. Teachers accordingly have specific representations about students from different ethnic backgrounds as well as about students from different socioeconomic levels. Thus, they have cognitive schemas—which are associated with more or less positive or negative attributes for these groups of students—on which they can fall back on if they have to form judgments or expectations.

The present work attempts to investigate whether differences exist in teachers' expectations and judgments depending on students' ethnic and social backgrounds. It aims to contribute to the existing literature in several ways. First, despite the large body of literature on the existence of stereotyped judgments in general, it remains an open question whether such judgments are a disadvantage for minority students or an advantage for majority students. Are teachers' expectations negatively or positively biased for stereotyped groups? As both potential biases may lead to different but important expectation effects, it is of high relevance to investigate positively or negatively biased teacher judgments. Second, as there is a strong linkage between immigrant background and socioeconomic level (e.g., Mehringer 2013), the present work attempts to analyze if a specific effect of ethnic stereotypes on teachers' judgments exists or if these disparities can be explained by stereotypes associated with the socioeconomic backgrounds of the students. Finally, aside from achievement expectations in a narrower sense, we additionally considered judgments of achievement-relevant characteristics (e.g., general ability, willingness to put in effort). Although these characteristics can be conceptualized as determining achievement—also in the view of teachers—little is known about potential biases when judging them.

To address these aspects, we conducted an experimental online study with German primary school teachers, who had to report expectations about students' future achievements and achievement-relevant characteristics. Using case vignettes, we experimentally varied students' ethnic and social backgrounds.

2 Expectation effects, teachers' judgments and students' diversity

2.1 Teachers' expectations and students' backgrounds

Teacher expectations are inferences that teachers form about a student's future achievement and behavior based on current student information (Good 1987). Expectations are generated on the basis of personal values and approaches and are based on the present knowledge about a person (e.g., Dusek et al. 1985). Some parts of this knowledge consist of characteristics that can be actually perceived, other factors are estimated by various visual characteristics, like age or mood, and further details are simply guessed. This occurs on the basis of personal experience with other people, the knowledge shared with society about social categories—which might be strongly linked to stereotypical assumptions—and mental probabilities (that do not necessarily need to be true, e.g. representativeness heuristic).

Assuming that teachers' interactions with their students are guided by teachers' values and attitudes, interactions are also mainly influenced by their expectations about students' achievements, behaviors and reactions to specific treatments. As long as expectations are accurate (e.g., adapted to actual abilities and possibilities, independent of social categories) there seem to be no negative consequences for students. But if teachers' expectations for future achievement and achievement-relevant characteristics are higher or lower than a given student's actual potential, it may lead to either a self-fulfilling prophecy (a wrong expectation leads to a specific behavior which causes the expected assumption; Jussim 1986; Jussim et al. 1996; Merton 1948) or, at least, a sustaining expectation effect (because of a teacher's expectation that a student's achievements will stay the same, changes in behavior and performance can go unnoticed; Cooper and Good 1983).

In both cases teachers' expectations can have effects on teacher-student-interactions (e.g., instructions, quality and quantity of feedback) whereby students draw conclusions about their own abilities and performances that influence their self-concept and motivation and, ultimately, actual performance (Good 1987; Good and Brophy 1997; McKown and Weinstein 2003). Thus, both types of expectation effects can influence students' learning processes as well as their school achievement and school success. Although the self-fulfilling prophecy effect is more popular, effects are often small. However, especially for stigmatized groups, they may have a substantial impact (Jussim and Harber 2005). The sustaining expectation effect is no less harmless than the self-fulfilling prophecy, and researchers assume that it occurs more often, but subtly (Good 1987; Good and Brophy 1997).

Rosenthal and Jacobson (1968a, 1968b) were the first to present empirical evidence of the influence of teacher expectations on actual students' achievements. Besides some methodical criticism, this study showed—especially for younger children—effects on their achievement as a result of teachers' positive expectations of their cognitive development. Rubie-Davis et al. (2006) showed for primary schools in New Zealand that teachers had lower expectations for the reading performance of students with a Maori background than for other ethnic groups regardless of students having the same reading skills at the beginning of the school year. Furthermore, students' gains in reading performance over one school year were less for those students for whom teachers had lower expectations at the beginning of the school year. A meta-analysis for the U.S. American context also showed higher expectations for Asian American and lower expectations for Latino and African American students compared with European American Students (Tenenbaum and Ruck 2007). Different expectations depending on a child's social background were first analyzed in an experimental study by Darley and Gross (1983). Participants watching an identical video of a girl's test performance varied in their expectations of ability depending on the experimental manipulation of beforehand presented information regarding a high versus low socioeconomic status of the girl. Compared to a low status, the assumed ability-level of the girl with high status was significantly higher in all three tested subjects (Liberal Arts, Reading, and Mathematics). Of course, different expectations could also be rooted in actual differences in students' performances, which could, however, already be the result

of teachers' expectations, making it quite difficult to clearly identify expectancy effects in correlational field studies. To overcome such confounding effects and to analyze the causal influences of social and ethnic backgrounds, experimental studies with controlled conditions are better suited—although they are associated with a certain loss of external validity.

As our knowledge is organized in an associative network (Anderson 1996), some social categories are linked with characteristics that might be of stereotype origin. Already in 1933, Katz and Braly showed that people associate different characteristics with different ethnic backgrounds and mainly link more positive characteristics to their own ethnic group and more negative characteristics to other ethnic groups. Glock and Karbach (2015) showed that student teachers had more negative implicit attitudes for minority students than for majority students. Casper et al. (2010) indicated that specific stereotype activation depends on the situational context. For example, the picture of a Turkish student near a place for prayers might activate a religious stereotype, but the picture of a Turkish student in a classroom might activate a different stereotype. Teachers know about international studies like PISA (Programme for International Student Assessment) or PIRLS (Progress in International Reading Literacy Study), and are aware of results that associate lower performance with an immigrant background and/or low socioeconomic status in comparison with majority students from high-income households. These findings and other personal convictions might lead to more negative associations for children with an immigrant background and/or low socioeconomic status in the school context. As people tend to link their own group with more positive attributes, teachers might also tend to associate more positive characteristics with students of their own ethnic group and have positive stereotypes for these students in the school context. It is important to note that popular results on achievement disparities associated with students' backgrounds may produce stereotypes that are correct for entire social groups (at least in the first instance, although disparities may be the effect of teachers' stereotypes). However, these stereotypes of course need not be true for the individual members of the social group.

Fiske and Neuberg (1990), however, showed in their continuum-model of impression formation that it is possible to switch from a category-based process to a controlled and thereby individual-based information process, but in order to do so cognitive capacities and motivation are needed. In not always seeing high relevance in everyday situations or not having the temporal or cognitive resources, we perceive people as representatives of social categories (Kahnemann 1973). Therefore, teachers can switch from controlled to automatic information processing, as their daily routine in teaching needs partly to be automated in order for them to act effectively in the classroom. These more category based processes are strongly linked to characteristics associated with social categories—which can be negative or positive. As it was already shown, (student) teachers have more negative associations for immigrant students (Glock and Karbach 2015), and this could lead to negative automated information processing, which would result in inaccurate and more negative judgments. On the other hand, it is also possible that positive stereotypes for high status students are the basis of category based information processing and therefore also lead to inaccurate, but positively biased judgments.

Beyond that, controlled information processing based on individual information about the student should lead to a higher probability of accurate judgments.

2.2 The case of Germany

As, in the context of ethnicity and social status, each country or region has its own specific conditions marked by a particular history as well as political and social country-specific situations, we cannot directly compare results from different regions. Comparable to children in other countries, children in Germany with an immigrant background and/or low socioeconomic status achieve lower competence levels in representative studies (e.g., Gebhardt et al. 2013; Müller and Ehmke 2013) and less frequently manage the transition to a higher secondary school (e.g., Gresch and Becker 2010). Specific to the German situation is that people of Turkish origin are the largest immigrant group (17.6%; Statistisches Bundesamt 2015). Compared with majority students or students with other immigrant backgrounds, like former USSR states, Turkish students achieve significantly lower. Furthermore, in contrast to other immigrant groups, the achievement gap exists not only for first generation immigrants, but also even for third generation immigrants and those with only one parent with Turkish roots (Gebhardt et al. 2013). As people with an immigrant background (especially a Turkish background) belong significantly more often to a low socioeconomic level (e.g., Mehringer 2013), it is not clear if effects of students' ethnicity could also (at least partly) be explained by their social background. Stereotypes for immigrant background as well as for a low socioeconomic status could be activated concurrently in this constellation. Teachers could automatically associate a low socioeconomic status when judging ethnic minority students. For Switzerland, Becker et al. (2013) showed that students were not disadvantaged because of their immigrant background but mainly because of their social level (see also Becker 2013).

In the German context, some studies have already analyzed teacher expectations and judgments based on students' ethnic and social backgrounds. For the pre-school context, Kratzmann (2013) showed lower expectations for immigrant children that partly can be explained by their social level. Additionally, some studies already indicated differences in primary and secondary teachers' judgments and expectations depending on students' backgrounds (e.g., Stahl 2007). Sprietsma (2013) showed the effects of students' ethnic background on teachers' grading in an experimental study with a variation of Turkish and German names on students' essays. In detail, primary school teachers evaluated the same essay worse when it was written by a student with a Turkish name. Furthermore, in the study of Sprietsma (2013), teachers less frequently recommended students with a Turkish name for a higher secondary school. In another experimental study, Glock and Krolak-Schwerdt (2013) showed that teacher judgments and expectations of a given student's performance in German and Mathematics depend on the student's nationality, but also on whether the information about the student confirms or disconfirms the expectation. Lorenz et al. (2016) showed that teachers had lower expectations for students with a Turkish background and/or low socioeconomic background in the subject of German, even when actual performance was

comparable with their majority classmates. Madon et al. (1998) showed under consideration of students' sex, social level and ethnicity that stereotypes were often quite accurate for these students. Their results indicate that judgments could not be exclusively explained by individual student characteristics. In sum, previous results indicate that teachers' judgments can be biased by students' backgrounds, and therefore are not always accurate.

2.3 Shortcomings of previous research

As we know from studies in the field of diagnostic competence, student teachers (novices) and experienced teachers (experts) differ in their judgment-processes, e.g. under conditions of inconsistent information, experts assessed students' character-istics more differentiated, and were less vulnerable to systematic judgmental distortion (van Ophuysen 2006). In general, stereotypes mainly influence automatic information processing which is more likely to occur when teaching experience is high because student teachers are usually not yet able to switch from controlled to automatic processes. As earlier studies mostly investigated student teachers, these results cannot be unrestrictedly transferred to experienced teachers. Besides this aspect, previous research also focused mainly on specific performance expectations, but neglected teachers' perceptions of achievement-relevant dispositions, like motivational aspects or ability. Furthermore, the strong connection between ethnic and social backgrounds (e.g., Kena et al. 2015) was often not considered, as usually both characteristics were explored separately. In one of the few exceptions, Ready and Chu (2015) showed for the U.S.-American case positively biased teachers' perceptions for high status preschool children in literacy development, and an underestimation of children with a low status or with a different primary language than English. As mentioned above, it is additionally an open question whether there is a disadvantage for minority students in teachers' achievement expectations in terms of negatively biased information processing, or a preference for majority students in terms of positively biased processing.

2.4 Research questions and hypotheses

Based on the theoretical assumptions and previous empirical research, this article investigates whether teachers have different achievement expectations, achievement aspirations, and judgments about achievement-relevant dispositions dependent on students' ethnic or social backgrounds. We considered teachers' perceptions of actual abilities, as they are a predictor of a student's academic performance (e.g., Harackiewicz et al. 2002), to play an important role in educational success and to be an essential basis for a teacher's orientation in planning lessons; therefore, accurate evaluation is needed. As a student's willingness to put in effort is a relevant part of achievement motivation and therefore an important predictor for school achieve-ment (e.g., Steinmayr and Spinath 2009), it also needs to be accurately evaluated by teachers. One basic task in primary school teachers' profession is to decide if students are qualified for a higher secondary school. As a result, teachers' decisions determine a student's future educational achievement. As mentioned above, the

activation of ethnic and social stereotypes by students' backgrounds may lead to biased teachers' judgments. According to Fiske and Neuberg (1990), automatic information processes rely on category based information which is associated with more or less positive or negative characteristics for each social category. So, potentially positive stereotypes for majority students could be activated and influence teachers' judgment to a positive bias. Correspondingly negative stereotypes for minority students may lead to a negative bias in judging.

To answer the question of whether teachers' judgments differ in relation to these background variables, we considered the following directional hypotheses:

H₁ Teachers' judgments of achievement-relevant characteristics, achievement expectations, and achievement aspirations for minority students are worse than those for majority students.

Students with an immigrant background and low socioeconomic status were defined as minority students, and students without an immigrant background and high socioeconomic status were defined as majority students. To answer questions of over- and underestimation, teachers' achievement expectations and achievement aspirations were focused.

H₂ Teachers' achievement expectations and aspirations are negatively biased for minority students.

H₃ Teachers' achievement expectations and aspirations are positively biased for majority students.

3 Method

3.1 Sample and procedure

We conducted an experimental online-study with $N = 237$ teachers from all over Germany. Teacher participation was voluntary and participants received a coupon for an online retailer as compensation for their effort. As in previous studies on teachers' judgments (e.g., Sprietsma 2013), case vignettes were used for experimentally manipulating students' backgrounds. We constructed case vignettes that represented the school report of three male primary school students. School reports were identical for all three boys, with only the first name experimentally varied between the case vignettes. Participants had to read one school report and, after doing so, answer questions about their achievement expectations and achievement aspirations for the student as well as their judgment about achievement-relevant characteristics, like the student's qualification for a higher secondary school, his general abilities and, as a motivational factor, his willingness to put in effort. As the case vignettes referred to fourth grade students, we recruited only primary school teachers (age: $M = 41.62$, $SD = 8.49$) with an average experience in teaching of $M = 15.34$ years ($SD = 8.37$). The sample consisted of 48.9% females.

3.2 Case vignettes

The case vignettes were comprised of a verbal and a numeric mid-year school report for three-fourth-graders. Although not only boys, but also girls with an immigrant background and/or low socioeconomic status may achieve lower performance in the educational system, we chose not to include girls' names in our study due to reasons of theoretical and empirical simplicity. Being that teachers may also hold a bundle of gender-specific stereotypes, a very extensive design would be necessary to additionally account for gender effects. A gender decision to focus on boys was made because the achievement gap and academic performance of boys with different ethnic and social backgrounds seem to be a major problem in Germany. Boys with Turkish roots compared with Turkish girls are underrepresented in higher secondary schools; furthermore, gender-differences in reading in favor of Turkish girls are much higher than gender differences for students without an immigrant background (Segeritz et al. 2010). The verbal part of the school report consisted of statements about the students' social, working and learning behaviors (phrases were selected from an official institute for schooling quality; Staatsinstitut für Schulqualität und Bildungsforschung München 2008). We balanced and rotated positive, neutral and negative phrases about students' behaviors to create an inconsistent picture. To avoid primacy- and recency-effects, neutral information was selected for the beginning and the end of the verbal school report. The numeric report also tried to present a heterogenic picture of the students, with German grades ranging from 1 (*very good*) to 3 (*satisfactory*) with an average of 2.11. In the main subjects of German, Mathematics, and Social Studies the grade point average was 2.33—a range were students can still attend a higher secondary school, but the decision is placed on the teacher.

3.3 Experimental variation of ethnic and social backgrounds

Stereotype activation resulted from variation of the students' names. For this reason, names had to clearly point out the ethnic as well as the social background. Utech (2011) analyzed naming statistics for Germany by parents' socioeconomic backgrounds. This served as the basis for our selection of the most frequent names in low and high social levels. Names that could be associated with particularly high or low cognitive abilities (Rudolph et al. 2007) were excluded. The Turkish name was selected from official birth statistics (Gesellschaft für deutsche Sprache 2012) that were assigned as typical Turkish names by student teachers (Tobisch 2013). Finally, three names of comparable length were selected and randomly assigned to participants: *Julius* (no immigrant background and high socioeconomic status; $n = 84$ teachers), *Justin* (no immigrant background and low socioeconomic status; $n = 77$) and *Murat* (immigrant background and low socioeconomic status; $n = 76$).

As a manipulation check we asked teachers at the end of the survey about their perception of the students' ethnic and social backgrounds. To check whether the teachers perceived the ethnic background in accordance with the manipulation, they were asked from which country the students' families came from in their opinion (Germany vs. other country); a significant and strong effect resulted ($\chi^2(2,$

$N = 237$) $= 228.06$, $p < .001$, Cramer's V $= .981$). For the perception of socioeconomic status, we used an item ("In your opinion, to which social level does the students' family belong to?") together with a 5-point rating scale ranging from 1 (*lower-level*) to 5 (*upper-level*). As theoretically expected, the social status of Murat ($M = 2.71$, $SD = 0.56$) and the social status of Justin ($M = 2.90$, $SD = 0.35$) were perceived as below average, and Julius ($M = 3.25$, $SD = 0.46$) was considered to be an above average status student ($F(2, 234) = 28.024$, $p < .001$, $\eta^2 = .19$).

3.4 Measures

We measured five central aspects that are important for a student's educational achievement as well as for a teacher's judgment about a student's future achievement. The expected future grades in the main subjects (German, Mathematics, and Social Sciences) were asked as well as the aspiration teachers had for a student's grades. General abilities, the willingness to put in effort, and teachers' opinions of qualification for a higher secondary school were also measured for each student.

Teachers' judgments about students' general abilities in school were measured with a 5-point bipolar scale and with four items ("I think the student's ability is …" from 1 [*low*] to 5 [*high*]; $\alpha = .88$; adapted from Dickhäuser et al. 2002). Teachers' opinions of a student's willingness to put in effort were measured with five items on a 5-point rating scale ranging from 1 (*I do not agree at all*) to 5 (*I totally agree*), ("The student will not give up, even when tasks are difficult and extensive"; $\alpha = .87$; adapted from Ramm et al. 2006). Teachers also had to report on a 5-point rating scale ranging from 1 (*not at all qualified*) to 5 (*very qualified*) on whether they thought the student was qualified for a higher secondary school ("Is the student qualified for a higher secondary school?"). With one item per subject we asked teachers what grade ranging from 1 (*very good*) to 6 (*insufficient*) they would expect the student to receive on the next test. Besides the measurement of actual achievement expectation, we also asked for teachers' aspiration levels for the student at hand. According to Finsterwald (2006), teachers had to report a grade from 1 (*very good*) to 6 (*insufficient*) that they would still be satisfied with in regard to the student's score on the next test in the three subjects. The grades were recoded so that high values indicated good grades.

Data were complete for all items except for students' ability scales (18.6% missing). Missing data were estimated and replaced with the expectation-maximation-algorithm (Graham 2009; Lüdtke et al. 2007).

3.5 Analyses

Besides descriptive statistics, we calculated multivariate analyses of variance for teachers' judgments of achievement-relevant characteristics, and for teachers' achievement expectations and achievement aspirations. Afterwards, we computed univariate analysis of variance for each construct. For detailed analysis of ethnic or social background effects, we conducted two a priori orthogonal Helmert contrasts (group means are compared in a previously settled combination in a way that allows

specific hypothesis testing; e.g., Field 2013) with the conditions low versus high socioeconomic status (contrast 1; *Murat* and *Justin* vs. *Julius*) and immigration background versus no immigration background (contrast 2; *Murat* vs. *Justin*). Finally, we calculated one sample t-tests for specific analyses of the accuracy of teachers' expectations.

4 Results

4.1 Descriptive results

Descriptive results (Table 1) show favorable teacher judgments for the average of all students. Bivariate correlations indicate an expected strong linkage between achievement-relevant characteristics, achievement expectations and achievement aspirations.

Table 1 Descriptive statistics and bivariate correlations of teachers' judgments of achievement-relevant characteristics as well as teachers' achievement expectations and aspirations

|	M (SD)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Judgments of achievement-relevant characteristics									
(1) General abilities	3.71 (0.47)								
(2) Willingness to put in effort	3.86 (0.66)	.76**							
(3) Qualification for higher secondary school	3.59 (0.71)	.59**	.49**						
Achievement expectations									
(4) German	5.05 (0.63)	.54**	.50**	.47**					
(5) Mathematics	4.95 (0.71)	.49**	.48**	.44**	.43**				
(6) Social studies	4.94 (0.66)	.51**	.52**	.50**	.65**	.53**			
Achievement aspirations									
(7) German	4.94 (0.74)	.58**	.43**	.57**	.51**	.53**	.51**		
(8) Mathematics	5.08 (0.87)	.57**	.47**	.51**	.48**	.53**	.45**	.45**	
(9) Social studies	4.98 (0.76)	.58**	.50**	.54**	.51**	.65**	.58**	.70**	.55**

$N = 237$

** $p < .01$

4.2 Teachers' judgments of achievement-relevant characteristics

Multivariate analysis of variance showed a main effect of students' backgrounds on teachers' judgments of achievement-relevant characteristics (Wilks $\Lambda = .859$, $F(6, 464) = 6.120$; $p < .001$, $\eta^2 = .07$), as illustrated in Fig. 1. Subsequent univariate analyses confirmed significant differences in teachers' judgments in dependence of students' backgrounds for all three achievement-relevant characteristics: General abilities ($F(2, 234) = 16.432$; $p < .001$, $\eta^2 = .12$), willingness to put in effort ($F(2, 234) = 7.540$; $p < .001$, $\eta^2 = .06$) and qualification for higher secondary school ($F(2, 234) = 10.550$; $p < .001$, $\eta^2 = .08$).

Results of analyzing the a priori contrasts for teachers' judgments of achievement-relevant characteristics, while simultaneously comparing low with high socioeconomic status (Contrast 1) and immigrants with non-immigrants (Contrast 2), can be found in Table 2. Findings revealed that differences in teachers' judgments could not be explained by either ethnic or social background alone—instead, both characteristics were important factors. Looking at teachers' judgments of general abilities, the student presented in the vignette with a high socioeconomic background condition was rated better than the student in the low status condition. As well, higher abilities were more attributed to the student in the non-immigrant condition than in the immigrant condition. The same pattern was identifiable for teachers' evaluations of willingness to put in effort. A non-immigrant background and high social status combination was greater associated with perceptions of the student as being more willing to try hard in school related tasks than in the low status and immigrant combination. Moreover, teachers rated the student in the vignette with an immigrant background as less qualified for a higher secondary school than the student in the vignette without immigrant background. Comparable results were shown for socioeconomic status. The student with a high status was
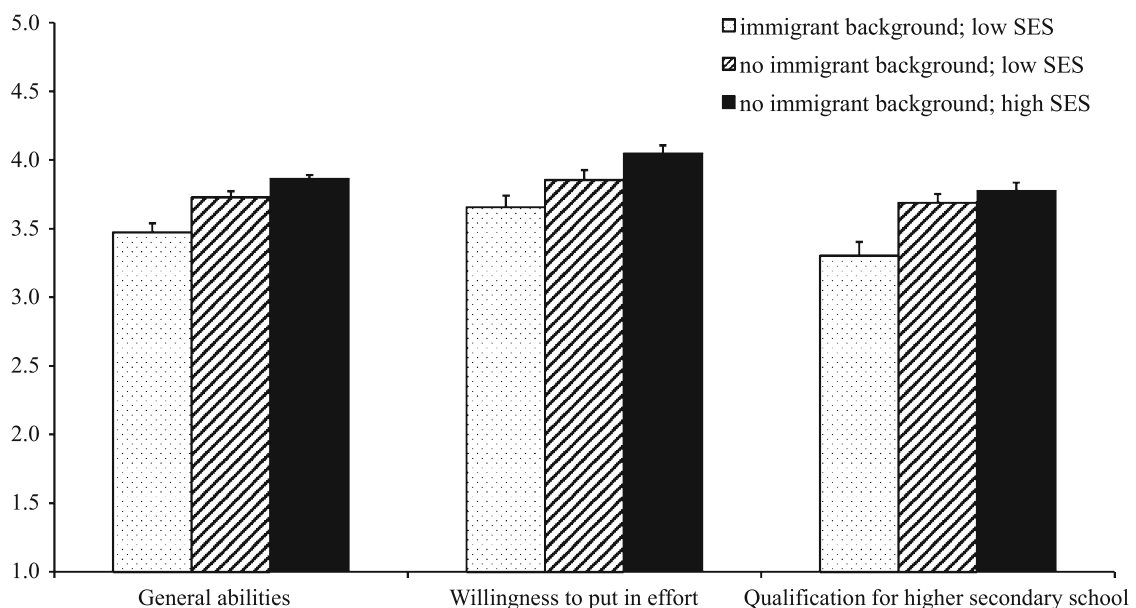
Fig. 1 Means and standard errors of teachers' judgments of achievement-relevant characteristics by students' backgrounds (*SES* socioeconomic status)

Table 2 Results of a priori contrast analyses for teachers' judgments of achievement-relevant characteristics as well as teachers' achievement expectations and achievement aspirations (SES = socioeconomic status)

|	Contrast	Contrast value	t	$r_{contrast}$
Achievement-relevant characteristics				
General abilities	Low versus high SES	0.266	4.461***	.28
	Immigrant versus non-immigrant	0.128	3.618***	.23
Willingness to put in effort	Low versus high SES	0.293	3.377***	.22
	Immigrant versus non-immigrant	0.100	1.931*	.13
Qualification for higher secondary school	Low versus high SES	0.278	3.000**	.19
	Immigrant versus non-immigrant	0.193	3.491***	.22
Achievement expectation				
German	Low versus high SES	0.155	1.832*	.12
	Immigrant versus non-immigrant	0.105	2.071*	.13
Mathematics	Low versus high SES	0.123	1.299	.08
	Immigrant versus non-immigrant	0.190	3.371***	.22
Social studies	Low versus high SES	0.313	3.593***	.23
	Immigrant versus non-immigrant	0.093	1.783*	.12
Achievement aspiration				
German	Low versus high SES	0.284	2.921**	.19
	Immigrant versus non-immigrant	0.165	2.846**	.18
Mathematics	Low versus high SES	0.479	4.311***	.27
	Immigrant versus non-immigrant	0.242	3.661***	.23
Social studies	Low versus high SES	0.218	2.177*	.14
	Immigrant versus non-immigrant	0.177	2.967**	.19

$N = 237$, $df = 234$

* $p < .05$; ** $p < .01$; *** $p < .001$

estimated more qualified for a higher secondary school than a student with a low status. Thus, both background variables seem to be relevant for teachers' judgments of achievement-relevant characteristics.

4.3 Achievement expectations and aspirations

We also found a main effect of students' backgrounds on teachers' achievement expectations and achievement aspirations on future tests in the three main subjects
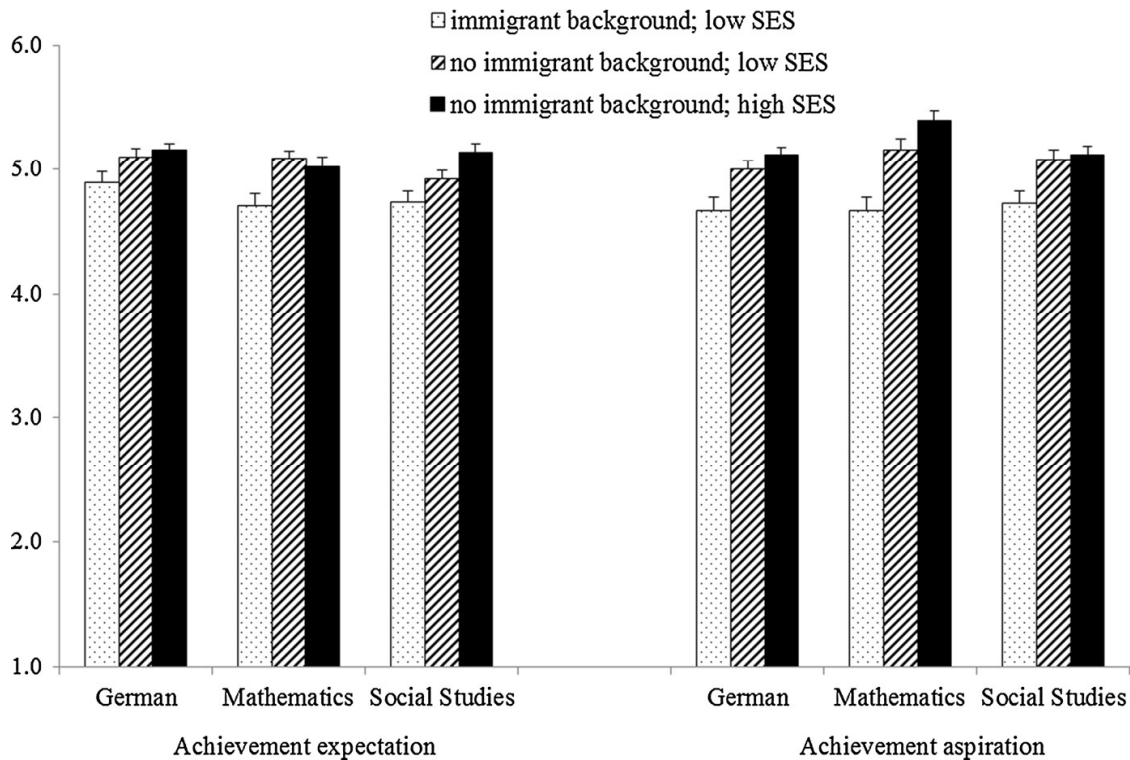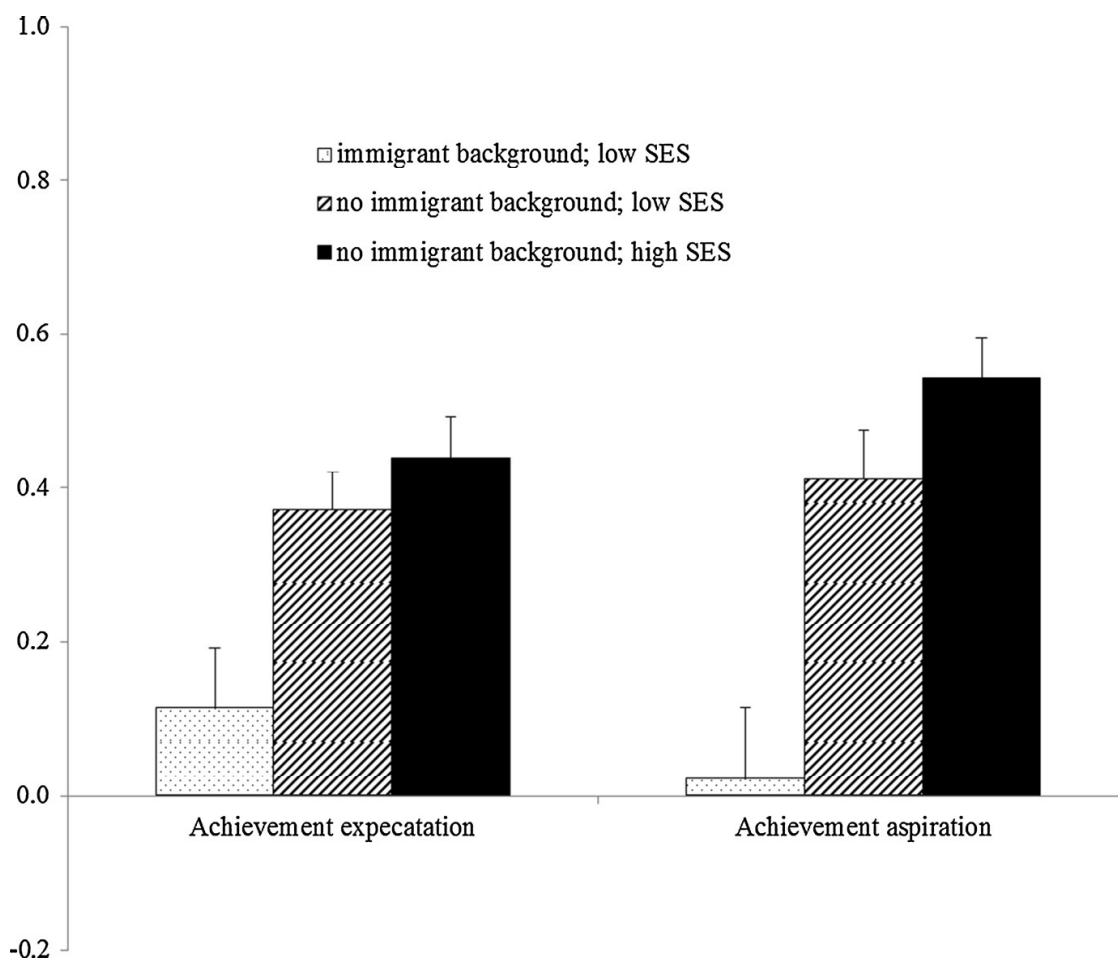
Fig. 2 Means and standard errors of teachers' achievement expectations and achievement aspirations for main subjects by students' backgrounds (*SES* socioeconomic status)

(Wilks $\Lambda = .820$, $F(12, 458) = 3.980$; $p < .001$, $\eta^2 = .09$). Respective means are presented in Fig. 2. Subsequent univariate analysis for each main subject showed that background variables had a significant influence on teachers' achievement expectations for German ($F(2, 234) = 3.808$; $p < .05$, $\eta^2 = .03$), Mathematics ($F(2, 234) = 6.509$; $p < .01$, $\eta^2 = .05$), and Social Sciences ($F(2, 234) = 8.019$; $p < .001$, $\eta^2 = .06$). The same patterns were evident when analyzing achievement aspirations for the three subjects (German: $F(2, 234) = 8.282$; $p < .001$, $\eta^2 = .07$; Mathematics: $F(2, 234) = 15.932$; $p < .001$, $\eta^2 = .12$; Social sciences: $F(2, 234) = 6.747$; $p < .001$, $\eta^2 = .06$).

In order to specifically examine whether differences in teachers' achievement expectations and achievement aspirations depend on either the students' ethnic backgrounds or their social backgrounds or both, we again analyzed the respective a priori contrasts (see Table 2). Once more, the results indicated that both ethnicity and social background are relevant factors in the formation of teachers' expectations for future grades in all main subjects. In detail, teachers had higher expectations and aspirations for the student with a high socioeconomic status in comparison with the two students in the low status vignettes across all subjects (one exception concerns the subject of Mathematics; see Table 2). Furthermore, comparing the two vignettes that present descriptions for students with low socioeconomic background revealed that teachers expected significantly better grades and hold higher aspirations for the students with no immigrant background than for the immigrant student.

In the final step, we tested if teachers' achievement expectations and aspirations were negatively biased for minority students, positively biased for majority

Fig. 3 Deviation (means and standard errors) of teachers' achievement expectations and achievement aspirations from the actual grade point average presented in the vignettes by students' backgrounds (*SES* socioeconomic status, *0* actual grade average point of 2.33 in the three main subjects)

students, or accurate. To this end, we calculated the differences between the actual grade point average and the teachers' achievement expectations and aspirations, averaged over the main subjects (Fig. 3).

For each of the three background configurations we conducted one sample t-tests to analyze if teachers' achievement expectations and achievement aspirations differed significantly from the actual grade point average (2.33) presented in the vignettes. For the combination of immigration background and low socioeconomic status, teachers' expectations $(t(75) = -1.417, p = .161)$ and aspirations $(t(75) = -0.201, p = .842)$ were quite accurate. However, for the combination of a non-immigrant background and a low socioeconomic status, teachers expected significantly better grades $(t(76) = -7.545, p < .001)$ and had higher aspirations $(t(76) = -6.315, p < .001)$ than the actual grade point average presented in the vignette. The difference between teachers' expectations and actual grade point average was also significantly higher for the non-immigrant and high socioeconomic status combination $(t(83) = -8.294, p < .001)$, and the same pattern was determined for teachers' achievement aspirations $(t(83) = -10.456, p < .001)$. In brief, teachers expected students with an immigrant background to achieve the same

performance as before, while they expected better achievements for students without an immigrant background in comparison to their actual achievements.

5 Discussion

The present work aimed to uncover background associated teacher expectations and judgments relating to primary school students. For this reason, we created fictive case vignettes of three-fourth graders' school reports with average grades as well as average learning and social behavior. In an experimental online study with experienced primary school teachers, we manipulated the students' first names so that they could be associated with specific ethnic and social backgrounds.

By using an experimental design that combined students' ethnic as well as social backgrounds, the present study allowed for conclusions to be made about the specific causal influences of both social categories on teachers' judgments and expectations. A further strength of the present study is that it not only focused on achievement expectation and achievement aspiration, but also on teachers' judgments about being qualified for a higher secondary school. Judgments about students' general abilities and students' willingness to put in effort were included as additional important factors for success in school. A final strength to be mentioned here is that the present study allowed for first indicators on whether teachers' judgments are negatively biased for minority students or positively biased for majority students.

As expected and in line with previous studies (e.g., Glock and Krolak-Schwerdt 2013; Rubie-Davies et al. 2006; Sprietsma 2013; Tenenbaum and Ruck 2007), results showed significant differences in teachers' judgments and expectations by ethnic and social backgrounds in all investigated areas. Teachers' expectations and judgments were lower for hypothetical students with a Turkish immigrant background and/or low socioeconomic status. According to our assumptions, teachers had the highest expectations under the condition of a high socioeconomic background and a non-immigrant background.

Contrary to our hypothesis, we could not find a negatively biased judgment for teachers' achievement expectations and achievement aspirations when comparing them with the actual grades from the case vignettes. As we could not find a significant difference between actual and expected grades under the condition of an immigrant background, it was determined that teachers seem to form accurate judgments for these students. With regard to the field of research on the inaccuracy of teachers' judgments (e.g., McElvany et al. 2009; Südkamp et al. 2012), our results imply clearly that different groups of students should be considered in future studies (cf. Kaiser et al. 2016). Teachers' judgments were more accurately in line with prior grades when they formed expectations for an immigrant student than when they had to judge a non-immigrant student with a low or high socioeconomic status.

One could argue that teachers should expect better grades for the future than the students had in the past. However, it is well documented that student achievement does not suddenly vary substantially from prior grades, and a negative trend to

worse grades is more usual than a trend to better grades (e.g., Hülür et al. 2011). Following this line of argumentation, we consider expectations of future grades that are in line with past grades as accurate. Consequently, we interpret teachers' achievement expectations and aspirations for students with an immigrant background as more accurate than those for students without an immigrant background. Additionally, our results can be seen as indicators of a positive bias for teachers' achievement expectations for majority students. In line with results of Ready and Chu (2015) teachers tend to overestimate high status students' future achievements. Although there was a difference between the high and low status German students, the main difference here was associated with ethnic background.

Considering the continuum-model of impression formation (Fiske and Neuberg 1990), results indicate different information processing dependent on a student's background. As the hypothetical student with a Turkish background was evaluated quite accurately, one could assume that teachers did not form their judgment based on category and automatic processes. Perhaps this is the case because they know about negative associations and are motivated to control their stereotyped associations. Positively biased judgments of the hypothetical German students (especially with high socioeconomic status) indicate that teachers have linked more positive stereotypes with this social category in the school context and may not be aware of the influence of positive associations on their judgments. Thereby they seem to not use individual-based and controlled information processing, but instead a positively biased category-based route of information processing. Certainly, further studies are needed to replicate the present findings and interpretations.

Of course, some limitations need to be mentioned. By designing an experimental study, we may conclude that students' backgrounds causally influenced teachers' decisions. However, the experimental design resulted in a loss of external validity. By using students' names as manipulation of the ethnic and social backgrounds, we could not include a case with a high status immigrant student. The strong linkage between a Turkish immigrant background and a low socioeconomic status (Mehringer 2013) may lead to combined associations for immigrant students. Therefore, future studies could use student information in case vignettes that explicitly point out a student's social level. It also has to be mentioned that teachers participated voluntarily in the study and therefore might be more motivated and engaged, which could have influenced their judgments. Furthermore, the results cannot be easily generalized to secondary schools, other ethnic backgrounds or gender differences. Moreover, we cannot tell how mental representation about ethnic and social categories are organized, or in what way they are mentally represented.

Besides some limitations, results indicate that different information processing and expectation formation occur for exactly the same student information depending on ethnic and social backgrounds. We suppose that each background is associated with specific stereotypes that lead to different mechanisms in terms of dual-process-theories of decision making. If motivation is high to control negative stereotypes, especially against ethnic minorities, information processing might be more controlled and less automatized. In knowing about the stereotypes, teachers may not want to act prejudiced and furthermore are probably motivated to accurately

judge their students and not to disadvantage some. As a consequence, they may judge quite accurately. In addition, positive stereotypes associated with a non-immigrant background and high socioeconomic status at first do not seem to be negative, so there is no need to process such information deeply if input is more or less interpreted consistently with prior attitudes and knowledge of this social category. At first sight, having high expectations for students does not seem to be negative in any way, and is not automatically connected with preference for some students. But in comparison with the different ethnic and/or social groups, having higher expectations for some students can be seen as positive discrimination or indirect negative discrimination for minority students.

According to theories about teacher expectation effects (e.g., Cooper and Good 1983; Jussim 1986; Jussim et al. 1996; Merton 1948), this may lead to two different, but nevertheless meaningful consequences. For students with a non-immigration background and high status, this might lead to a self-fulfilling prophecy. The expected positive development can result in more challenges and support from teachers and in the end lead to actual higher performance and improved behavior. As expectations for immigrant students seem to be more orientated on actual previous performance, teachers do not form higher expectations that could challenge students and thereby improve skills and performance. For this reason, it is an even bigger challenge for immigrant students to show their abilities and receive adequate encouragement and support. Future studies need to focus on whether this can actually be explained by different information processes, and on the question of whether accurate expectations for minority students lead to a sustaining expectation effect. If the present interpretations are applied to real classrooms, they provide a potential explanatory factor for the still existing gap between students with different ethnic and social backgrounds.

References

American Psychological Association (2012). *Ethnic and racial disparities in education: Psychology's contributions to understanding and reducing disparities.* Retrieved from http://www.apa.org/ed/resources/racial-disparities.aspx.

Anderson, J. R. (1996). *Kognitive psychologie [Cognitive psychology]* (2nd ed.). Heidelberg: Spektrum Akademischer Verlag.

Aud, S., Husser, W., Planty, M., Snyder, T., Bianco, K., Fox, M., et al. (2010). *The condition of education 2010 (NCES 2010-028): National center for education statistics, institute of education sciences.* Washington, DC: U.S. Department of Education.

Becker, R. (2013). Bildungsungleichheit und Gerechtigkeit in der Schweiz [Educational inequality and justice in Switzerland]. *Schweizerische Zeitschrift für Bildungswissenschaften, 35*(3), 405–413.

Becker, R., Jäpel, F., & Beck, M. (2013). Diskriminierung durch Lehrpersonen oder herkunftsbedingte Nachteile von Migranten im Deutschschweizer Schulsystem. [Ethnic penalties or discrimination by teachers towards migrants in the school system in the German part of Switzerland?]. *Swiss Journal of Sociology, 39,* 517–549.

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind.* New York, NY: Russell Sage.

Brimi, H. M. (2011). Reliability of grading high school work in English. *Practical Assessment, Research & Evaluation, 16*(17), 1–12. Retrieved from http://pareonline.net/getvn.asp?v=16&n=17.

Casper, C., Rothermund, K., & Wentura, D. (2010). Automatic stereotype activation is context dependent. *Social Psychology, 41,* 131–136. doi:10.1027/1864-9335/a000019.

Cooper, H., & Good, T. (1983). *Pygmalion grows up: Studies in the expectation communication process.* White Plains, NY: Longman.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44,* 20–33. doi:10.1037/0022-3514.44.1.20.

Dee, T. S. (2005). A teacher like me. Does race, ethnicity, or gender matter? *The American Economic Review, 95,* 158–165. Retrived from: http://www.jstor.org/stable/4132809.

Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept: Konstruktion und Überprüfung eines neuen Instruments [The academic self-concept scales: Construction and evaluation of a new instrument]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 23,* 393–405. doi:10.1024//0170-1789.23.4.393.

Dusek, J. B., Hall, V. C., & Meyer, W. J. (Eds.). (1985). *Teacher expectancies.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Esser, H. (2006). *Migration, Sprache und Integration: AKI-Forschungsbilanz 4 [Migration, language and integration: AKI-research results].* Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock'n'roll.* Los Angeles, CA: Sage.

Finsterwald, M. (2006). *Motivation und Schulübergang: Bedingungsfaktoren eines erfolgreichen Übergangs auf weiterführende Schulen [Motivation and school transition: Factors for a successful transition to secondary school].* Berlin: Logos.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social cognition* (pp. 231–254). New York, NY: Guilford Press.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). San Diego, CA: Academic Press.

Fuhr, G. (2012). *Armutsgefährdung von Menschen mit Migrationshintergrund: Ergebnisse des Mikrozensus 2010 [Risk of poverty of people with immigrant background: Results of micro-census 2010].* Wiesbaden: Statistisches Bundesamt, Wirtschaft und Statistik.

Gebhardt, M., Rauch, D., Mang, J., Sälzer, C., & Stanat, P. (2013). Mathematische Kompetenzen von Schülerinnen und Schülern mit Zuwanderungshintergrund [Mathematical competences of students with immigrant background]. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Eds.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (pp. 275–308). Münster: Waxmann.

Gesellschaft für deutsche Sprache. (2012). *Vornamenstatistiken Bayern 2010–2012 [First names statistics for Bavaria 2010–2012].* Wiesbaden: GfdS. (unveröffentlicht).

Glock, S., & Karbach, J. (2015). Preservice teachers' implicit attitudes toward racial minority students: Evidence from three implicit measures. *Studies in Educational Evaluation, 45,* 55–61. doi:10.1016/j.stueduc.2015.03.006.

Glock, S., & Krolak-Schwerdt, S. (2013). Does nationality matter? The impact of stereotypical expectations on student teachers' judgments. *Social Psychology of Education, 16,* 111–127. doi:10.1007/s11218-012-9197-z.

Glock, S., Krolak-Schwerdt, S., Klapproth, F., & Böhmer, M. (2013). Beyond judgment bias: How students' ethnicity and academic profile consistency influence teachers' tracking judgments. *Social Psychology of Education, 16,* 555–573. doi:10.1007/s11218-013-9227-5.

Good, T. L. (1987). Two decades of research on teacher expectations: Findings and future directions. *Journal of Teacher Education, 38,* 32–47. doi:10.1177/002248718703800406.

Good, T. L., & Brophy, J. (1997). *Looking in classrooms.* London: Longman.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. doi:10.1146/annurev.psych.58.110405.085530.

Gresch, C., & Becker, M. (2010). Sozial- und leistungsbedingte Disparitäten im Übergangsverhalten bei türkischstämmigen Kindern und Kindern aus (Spät-)Aussiedlerfamilien [Social and achievement conditioned disparities in transition behavior of children with Turkish immigration background and (late) repatriates' children]. In BMBF (Ed.), *Der Übergang von der Grundschule in die*

weiterführende Schule: Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten (pp. 181–200). Bonn: BMBF.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94,* 562–575. doi:10.1037//0022-0663.94.3.562.

Hauser, R. M., Simmons, S. J., & Pager, D. I. (2000). *High school dropout, race-ethnicity and social background from the 1970s to 1990s.* New York, NY: Russel Sage Foundation.

Hülür, G., Wilhelm, O., & Robitzsch, A. (2011). Multivariate Veränderungsmodelle für Schulnoten und Schülerleistungen in Deutsch und Mathematik [Multivariate change models for student achievement and school grades in German and mathematics]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 43,* 173–185. doi:10.1026/0049-8637/a000051.

Ingenkamp, K. (1989). *Die Fragwürdigkeit der Zensurengebung: Texte und Untersuchungsberichte [Dubiousness of grading. Texts and investigation reports]* (8th ed.). Weinheim: Beltz.

Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review, 93,* 429–445. doi:10.1037/0033-295X.93.4.429.

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

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9,* 131–155. doi:10.1207/s15327957pspr0902_3.

Kahnemann, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kaiser, J., Südkamp, A., & Möller, J. (2016). The effects of student characteristics on teachers' judgment accuracy: Disentangling ethnicity, minority status, and achievement. *Journal of Educational Psychology (published online).* doi:10.1037/edu0000156.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology, 28,* 280–290. doi:10.1037/h0074049.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X. Rathbun, A. Zhang, J., & Wilkinson-Flicker, S. (2015). *The condition of education 2015 (NCES 2015-144).* Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubs2015/2015144.pdf.

Kluczniok, K., Lehrl, S., Kuger, S., & Rossbach, H.-G. (2013). Quality of the home learning environment during preschool age: Domains and contextual conditions. *European Early Childhood Education Research Journal, 21,* 420–438. doi:10.1080/1350293X.2013.814356.

Kratzmann, J. (2013). Migrationsgekoppelte Ungleichheit durch niedrige Erwartungen im Kindergarten? Fähigkeitseinschätzungen und Prognosen durch Erzieherinnen [Are educational inequalities of children from immigrant backgrounds reinforced by lower expecations? Ratings and predictions of preschool teachers]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 65,* 73–99. doi:10.1007/s11577-013-0194-9.

Kristen, C. (2002). Hauptschule, Realschule oder Gymnasium. Ethnische Unterschiede am ersten Bildungsübergang [Hauptschule, realschule or gymnasium? Ethnic differences at the transition from primary to secondary schooling]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 54,* 534–552. doi:10.1007/s11577-002-0073-2.

Krohne, J., Meier, U., & Tillmann, K.-J. (2004). Sitzenbleiben, Geschlecht und Migration. Klassen-wiederholung im Spiegel der PISA-Daten [Non-promotion, gender, and migration—The repetition of school years as reflected in the PISA-study]. *Zeitschrift für Pädagogik, 50,* 373–391. Retrieved from http://nbn-resolving.de/urn:nbn:de:0111-opus-48161.

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,* 89–111. doi:10.1007/s11577-015-0352-3.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen [Handling of missing data in psychological research: Problems and solutions]. *Psychologische Rundschau, 58,* 103–117. doi:10.1026/0033-3042.58.2.103.

Madon, S., Jussim, L., Keiper, S., Eccles, J., Smith, A., & Palumbo, P. (1998). The accuracy and power of sex, social class, and ethnic stereotypes: A naturalistic study in person perception. *Personality and Social Psychology Bulletin, 24*(12), 1304–1318. doi:10.1177/01461672982412005.

Martin, M. O., Mullis, I. V. S., & Foy, P. (2012). *TIMSS 2011 International results in Science*. Chestnut Hill, MA: IEA.

McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., et al. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern [Teachers' diagnostic skills to judge student performance and task difficulty when learning materials include instructional pictures]. *Zeitschrift für Pädagogische Psychologie, 23*(2–3), 223–235. doi:10.1024/1010-0652.23.34.223.

McKown, C., & Weinstein, R. S. (2003). The development and consequence of stereotype consciousness in middle childhood. *Child Development, 74,* 498–515. doi:10.1111/1467-8624.7402012.

Mehringer, V. (2013). *Weichenstellung in der Grundschule: Sozial-Integration von Kindern mit Migrationshintergrund [Decisions in primary schools: Social integration of children with immigrant backgrounds]*. Münster: Waxmann.

Merton, R. K. (1948). *The self-fulfilling prophecy. The Antioch Review, 8,* 193–210. doi:10.2307/4609267.

Müller, K., & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung [Social background as condition for competence development]. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Eds.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (pp. 245–274). Münster: Waxmann.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International results in reading*. Chestnut Hill, MA: IEA.

OECD. (2012). Equity and quality in education. Supporting disadvantaged students and schools. *OECD*. doi:10.1787/9789264130852-en.

OECD. (2016). *Low-performing students: Why they fall behind and how to help them succeed*. Paris: PISA OECD Publishing.

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. *Journal of Educational Research, 106,* 64–76. doi:10.1080/00220671.2012.667010.

Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., et al. (Eds.). (2006). *PISA 2003. Dokumentation der Erhebungsinstrumente [PISA 2003. Documentation of measuring instruments]*. Münster: Waxmann.

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

Rosenthal, R., & Jacobson, L. (1968a). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Holt, Rinehart and Winston.

Rosenthal, R., & Jacobson, L. (1968b). Pygmalion in the classroom. *The Urban Review, 3,* 17–20. doi:10.1007/BF02322211.

Rubie-Davis, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology, 76,* 429–444. doi:10.1348/000709905x53589.

Rudolph, U., Böhm, R., & Lummer, M. (2007). Ein Vorname sagt mehr als 1000 Worte: Die soziale Wahrnehmung von Vornamen [A name says more than a thousand words: The social perception of first names]. *Zeitschrift für Sozialpsychologie, 38,* 17–31. doi:10.1024/0044-3514.38.1.17.

Schwippert, K., Wendt, H., & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund [Reading competences of students with immigrant background]. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Eds.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschlang im internationalen Vergleich* (pp. 191–208). Münster: Waxmann.

Segeritz, M., Stanat, P., & Walter, O. (2010). Muster des schulischen Erfolgs von Mädchen und Jungen mit Migrationshintergrund [Patterns of educational achievement of girls and boys with immigrant background]. In C. Allemann-Ghiond, P. Stanat, K. Göbel, & C. Röhner (Eds.), *Migration, Identität, Sprache und Bildungserfolg* (pp. 165–186). Weinheim: Beltz. (Zeitschrift für Pädagogik, Beiheft 55).

Sprietsma, M. (2013). Discrimination in grading: Experimental evidence from primary school teachers. *Empirical Economics, 45*(1), 523–538. doi:10.1007/s00181-012-0609-x.

Staatsinstitut für Schulqualität und Bildungsforschung München. (2008). *Formulierungshilfen zu den Grundschul-Zeugnissen: Formulierungshilfen und Textbausteine zum neuen Zeugnis [Help for*

formulation of primary school reports: Help for formulation and text blocks for the new school report]. München: Staatsinstitut für Schulqualität und Bildungsforschung.

Stahl, N. (2007). Schülerwahrnehmung und -beurteilung durch Lehrkräfte [Student perception and evaluation by teachers']. In H. Ditton (Ed.), *Kompetenzaufbau und Laufbahnen im Schulsystem: Ergebnisse einer Längsschnittuntersuchung an Grundschulen* (pp. 171–198). Münster: Waxmann.

Statistisches Bundesamt. (2015). *Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2014 [Population and employment. Results of micro-census 2014]*. Wiesbaden: Statistisches Bundesamt.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19,* 80–90. doi:10.1016/j.lindif.2008.05.004.

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

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology, 99,* 253–273. doi:10.1037/0022-0663.99.2.253.

Tobisch, A. (2013). Ethnische Stereotype und diagnostische Kompetenz von Lehramtsstudierenden. [Ethnic stereotypes and diagnostic competence of student teachers] (Unveröffentlichte Masterarbeit). Augsburg, Universität Augsburg.

Utech, U. (2011). *Rufname und soziale Herkunft: Studien zur schichtspezifischen Vornamensvergabe in Deutschland [First name and social background: Studies of social status specific naming in Germany]*. Hildesheim: Olms.

van Ophuysen, S. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahnempfehlung [Comparison of diagnostic decisions between novices and experts: The example of school career recommendation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38,* 154–161. doi:10.1026/0049-8637.38.4.154.

Wendt, H., Stubbe, T. C., & Schwippert, K. (2012). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern [Social background and students' reading competence]. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Eds.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 175–190). Münster: Waxmann.

Anita Tobisch is a researcher affiliated with the Department of Psychology at the University of Augsburg, Germany. Her research interests include diversity in educational contexts as well as teachers' judgments and diagnostic competence.

Markus Dresel is full professor affiliated with the Department of Psychology at the University of Augsburg, Germany. His research focuses on motivation and self-regulated learning in educational contexts as well as the relation between students' gender, ethnic and social backgrounds with educational success.