

University of Groningen

Accurate, inaccurate, or biased teacher expectations

Timmermans, Anneke C.; Kuyper, Hans; van der Werf, Greetje

Published in:
British journal of educational psychology

DOI:
[10.1111/bjep.12087](https://doi.org/10.1111/bjep.12087)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Timmermans, A. C., Kuyper, H., & van der Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *British journal of educational psychology*, 85(4), 459-478. <https://doi.org/10.1111/bjep.12087>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education?

Anneke C. Timmermans*, Hans Kuyper and Greetje van der Werf
GION Education/Research, University of Groningen, The Netherlands

Background. In several tracked educational systems, realizing optimal placements in classes in the first year of secondary education depends on the accuracy of teacher expectations.

Aims. The aim of this study was to investigate between-teacher differences in their expectations regarding the academic aptitude of their students.

Sample. The sample consisted of 500 teachers (classes) who provided their expectations of 7,550 students in the final grade of Dutch primary education.

Methods. We analysed the extent to which teachers differed in their expectations and in what contexts their expectations were biased, using multilevel random slope models.

Results. Multilevel analysis showed teacher expectation bias to be related to gender and socio-ethnic background of students. The differences among teachers in expectations for Turkish, Moroccan, and other foreign students with low-educated parents were larger than the average teacher expectation bias for these groups in the sample. When student characteristics were controlled for, we found that the teachers in our sample had higher expectations for students in high-performing classes or classes with only a small proportion of students from low-SES families. Teacher expectation bias for demographic groups, however, was found to be independent of the class population.

Conclusions. The results indicate the importance of the teacher and the necessity of using multilevel models when investigating the complex nature of between-teacher differences in expectations of student performance.

The transition from primary to secondary education is a crucial moment in the educational careers of students, where unequal opportunities may arise, especially in tracked secondary school systems, such as those of the Netherlands, Belgium, Germany, Hong Kong, and Singapore. In the Dutch system, the placement of students in the various tracks in the first year of secondary education mainly depends on their scores on a standardized primary school leaving test and the teacher's track recommendations (Driessen, 2005). The track recommendation can be considered as the expression of the teacher's expectations for the student's future performance during secondary education (De Boer, Bosker, & Van der Werf, 2010; Inspectorate of Education, 2007). Therefore, the accuracy of teachers' expectations is important in realizing optimal placements in classes in the first year of secondary education (Helwig, Anderson, & Tindal, 2001) and partly determines

*Correspondence should be addressed to Anneke C. Timmermans, GION Educational/Research, University of Groningen, Grote Rozenstraat 3, 9712 TG, Groningen, The Netherlands (email: a.c.timmermans@rug.nl).

students' opportunities during secondary and further education (Südkamp, Kaiser, & Möller, 2012).

In the Netherlands, an ongoing (political) debate has taken place in recent decades concerning the accuracy of the recommendations and the possible consequences of inaccurate recommendations for the educational careers of students. Central to this debate are (1) equity in the educational opportunities of ethnic minority groups and students with low-educated parents and (2) the question whether possible bias in recommendations for these groups contributes to the existing achievement gap (De Boer, Bosker, & Van der Werf, 2007; Driessen, 2005, 2011; Dronkers, van Erp, Robijns, & Roeleveld, 1998; Timmermans, Kuyper, & van der Werf, 2013). Although this debate initially focused on ethnic minority groups in general, attention gradually shifted to the educational opportunities of students from Turkish and Moroccan backgrounds, as they constitute the largest minority groups in the Netherlands. Although there is substantial variation, children of Turkish and Moroccan families typically enter primary education with language shortcomings and at the end of primary education, they show a 2-year language and a 6-month mathematics delay (Social and Cultural Planning Office, 2005).

The current study focuses on the transition from primary to secondary education as a crucial point in students' educational careers, and aims at investigating (1) whether teachers generally are less accurate in their expectations for subgroups of students based on their demographic characteristics and after their performance is taken into account, (2) whether there are between-teacher differences in their expectations for students with similar performance records, and (3) whether teachers' expectations are associated with class composition.

Teacher expectations: Accuracy, inaccuracy, and bias

The term *teacher expectations* refers to inferences made by teachers with respect to students' potential to achieve (Riley & Ungerleider, 2012). Research in the field of teacher expectations began with the experimental study *Pygmalion in the Classroom* by Rosenthal and Jacobsen (1968), which demonstrated that when teachers had high expectations of students, these students tended to confirm the expectations, a phenomenon known as 'self-fulfilling prophecy' (Merton, 1948). The magnitude of self-fulfilling prophecy effects in naturalistic studies of teacher expectations generally seems to be relatively small ($r = .10$ to $r = .20$). However, especially low achievers (Madon, Jussim, & Eccles, 1997) and students from low-income families and minority groups (Hinnant, O'Brien, & Ghazarian, 2009; Jussim, Eccles, & Madon, 1996; McKown & Weinstein, 2002, 2008; Sorhagen, 2013) are more susceptible to self-fulfilling prophecy effects. Because self-fulfilling prophecy effects can only occur when the teacher's expectation was initially inaccurate (Madon, Willard, Gyll, & Scherr, 2011; Merton, 1948), another body of literature relates to whether there are particular subgroups of students for whom teachers are less accurate in their expectations.

In many of these studies, accuracy is defined as the teacher's expectation in relation to the student's previous performance (Cooper, Findley, & Good, 1982; De Boer *et al.*, 2010; Hinnant *et al.*, 2009). A commonly used measure of accuracy is the discrepancy index (residual scores) as derived from a regression model in which the expectations of teachers are regressed on the previous performance of students. The closer to zero a student's residual score is, the more accurate the expectation of the teacher was (Hinnant *et al.*, 2009; Madon *et al.*, 1997). There is a general consensus among educational psychologists that most teachers have fairly accurate expectations of their students (Brophy, 1983;

Jussim & Eccles, 1992; Jussim & Harber, 2005; Südkamp *et al.*, 2012). However, teachers' expectations can be biased in two ways: Generally or specifically. Bias arises if the teacher's expectations are systematically too high or too low for most students in their class (general bias) or when teachers systematically expect too much or too little from specific subgroups of students (specific bias). The latter is sometimes referred to as differential teacher expectations (Van den Bergh, Denessen, Hornstra, Voeten, & Holland, 2010). Biased teacher expectations, however, only include the *systematic* differences between teacher expectations and the students' prior achievement. Inaccuracy, on the other hand, involves *all* differences between teacher expectations and the students' prior achievement, including random deviations. Therefore, biased teacher expectations are inaccurate, but inaccurate teacher expectations are not necessarily biased (Ready & Wright, 2011).

Evidence of teacher expectations bias

A number of non-experimental or naturalistic studies have looked at specific teacher expectation bias in relation to students' demographic characteristics, such as socio-economic status or minority status, thereby partly investigating on which information teachers base their expectations (Dusek & Joseph, 1982). In the Dutch context, several studies have investigated whether inaccurate recommendations at the end of primary education were more prevalent among students from ethnic minority groups or students from low socio-economic backgrounds. In general, it was concluded that there appeared to be no structural bias towards or against students from ethnic minority groups (Driessen, 2005, 2011; Driessen & Smeets, 2007; Roeleveld, Driessen, Ledoux, Cuppen, & Meijer, 2011); however, bias in teacher expectations against students from low socio-economic backgrounds was consistently found (De Boer *et al.*, 2010; Timmermans *et al.*, 2013). In the Dutch context, the prevalence of unduly low recommendations for Dutch students from low socio-economic backgrounds is particularly disquieting (Claassen & Mulders, 2003). All in all, the prior performance of students appeared to be the dominant factor for determining recommendations (Luyten & Bosker, 2004).

In the international context, studies of teacher expectation bias have resulted in inconsistent findings concerning bias against or in favour of particular subgroups of students (McKown & Weinstein, 2008; Ready & Wright, 2011). On the one hand, Jussim and Harber (2005) concluded in their review that differences in teacher expectations with respect to stigmatized demographic subgroups closely corresponded to differences in these groups' academic performance. In other words, they found no evidence for teacher expectation bias for stigmatized student subgroups. On the other hand, a meta-analysis showed that teacher expectations (and subsequent teacher behaviour) varied depending on the students' ethnic background after the performance of students was taken into account (Tenenbaum & Ruck, 2007). Other recent findings suggest that, on average, teachers tend to have lower expectations for the future academic performance of minority students and students from less affluent families (Glock & Krolak-Schwerdt, 2013; Glock, Krolak-Schwerdt, Klapproth, & Böhmer, 2013; Rubie-Davies, Hattie, & Hamilton, 2006; Speybroeck *et al.*, 2012).

Common to most of these studies is that 'overall' effects were investigated by taking all teachers together. Pooling data across teachers in reporting summary statistics of (in) accuracy and bias, however, obstructs the possibility of recognizing possibly important individual differences between teachers (Hoge & Coladarci, 1989). The few studies that have looked at between-teacher differences in *general* bias suffered from several

methodological problems, such as inadequate sampling (De Boer *et al.*, 2010) or reporting on differences among teachers before including student performance variables in the models (McKown & Weinstein, 2008; Van den Bergh *et al.*, 2010). In these studies, 17% of the variance in teacher expectations reported was associated with the class level (McKown & Weinstein, 2008) and 13% with the teacher level (Van den Bergh *et al.*, 2010). In the study by De Boer *et al.* (2010), about 26% of the variance in teacher expectations was found at the school level, after several student characteristics were controlled for, including student performance, intelligence, and performance motivation. The relatively large class- and teacher-level variances indicate that students who have similar records of prior achievement and background characteristics may receive very different expectations or recommendations from class to class or from teacher to teacher.

In addition to differences in *general* bias between teachers, teachers may also differ in their bias towards or against *specific* groups of students. The characteristics of the student population of their class may affect the degree to which teachers use student background as a guideline for their expectations (McKown & Weinstein, 2008; Ready & Wright, 2011). This may affect the teachers' accuracy and lead to individual differences in specific bias. McKown and Weinstein (2008) propose that when classes are 'mixed' (e.g., containing multiple groups of ethnically similar students), teachers may base their expectations on general group characteristics rather than on individual characteristics, due to directly perceived contrasts. In this context, teachers may be more likely to apply stereotypes about ethnic groups to individuals rather than form expectations about individual students based on their individual characteristics. In their study, McKown and Weinstein (2008) found a significant interaction between classroom diversity (number of ethnic groups) and teacher expectation bias related to students' ethnicity. They concluded that bias against specific ethnic groups was larger in classes with many ethnic groups than in class that consisted of fewer ethnic groups.

The current study

In this study, we aimed to investigate (1) whether teachers are generally less accurate in their expectations for subgroups of students based on demographic characteristics, (2) whether there are between-teacher differences in their expectations for students with similar performance records, and (3) whether the teachers' expectations are associated with the composition of the student population in their classes. We tested the following hypotheses in a large sample of 7,550 students in Dutch primary education. With respect to teacher expectation accuracy and bias for student subgroups, we formulated two hypotheses based on the current knowledge base. First, we expected the teacher recommendations to correlate strongly with the prior performance of students (H1), as teachers have been found to be relatively accurate in general (Brophy, 1983; Jussim & Eccles, 1992; Jussim & Harber, 2005; Südkamp *et al.*, 2012) and performance is the key determinant of the track recommendations at the end of Dutch primary education (Luyten & Bosker, 2004). Second (H2), we expected specific bias against students from low socio-economic backgrounds, but no differences between Dutch students and students from ethnic minority groups (De Boer *et al.*, 2010; Driessen, 2011; Driessen & Smeets, 2007; Roeleveld *et al.*, 2011; Timmermans *et al.*, 2013). With respect to differences in expectations between teachers, it was possible to formulate only tentative hypotheses, because many of the above-mentioned studies investigated 'overall' effects by taking all teachers together. Based on the findings of the three studies that looked into class and school differences in expectations (De Boer *et al.*, 2010; McKown & Weinstein, 2008;

Van den Bergh *et al.*, 2010), we expected teachers to differ in general expectation bias (i.e., systematically expecting too much or too little of all students in their class; H3). Finally, based on the assumption that the student populations of classes differ and that the characteristics of the student populations of classes may affect the degree to which teachers use student background as the guideline for their expectations (McKown & Weinstein, 2008; Ready & Wright, 2011), we predicted that there would be differences among teachers in specific bias towards or against student subgroups (H4). However, this final expectation cannot yet be grounded in empirical findings.

Method

Sample

The analyses were conducted on an existing set of data that were collected in the final grade (grade 6, age approximately 12 years) of primary education in 2004/2005 (Driessen, van Langen, & Vierke, 2006). The data consisted of a sample of 420 primary schools that were representative of Dutch primary schools, supplemented with a sample of 180 schools with relatively high proportions of minority students, to ensure that sufficient variation in school composition was available in the total sample.

The original sample contained records of 11,057 students in 630 classes for whom track recommendations were available. For 3,507 students in 130 classes, the values of one or more predictor variables were missing. These students were excluded from the analysis. This resulted in a data set containing the records of 7,550 students in 500 classes.¹ Of the 130 classes that were excluded from the analyses due to missing values, most were excluded because values on the school leaver's test or the tests from the monitoring systems were missing for all students. Comparing the information of the students included in the analysis with that of students who were omitted revealed some signs of attrition bias. The teacher recommendations were higher for the students included in the analysis (effect size: $d = .20$, $t = 10.26$; $df = 6296.2$; $p < .001$). Furthermore, the average score on the school leaver's test was higher for the students included in the analysis ($d = .25$, $t = 8.05$; $df = 1777.8$; $p < .001$). Similarly, the students included in the analyses performed higher on average in the three performance tests (language: $d = .13$, $t = 6.05$; $df = 5384.4$; $p < .001$; mathematics: $d = .09$, $t = 3.78$; $df = 4501.0$; $p < .001$; reading comprehension: $d = .12$, $t = 5.16$; $df = 4832.1$; $p < .001$).

Variables and instruments

Of focal interest in this study was the variable 'teacher expectations'. Furthermore, four student performance variables were available: The school leaver's test and three additional tests for language, mathematics, and reading comprehension during the final year of primary education. Finally, gender and socio-ethnic background were used to test whether or not the teachers were specifically biased in their expectations. These variables are described below; their descriptive statistics are provided in Table 1.

¹ Some classes may have had two part-time teachers. Some expectations included in this study might have been shared ones.

Table 1. Descriptive characteristics of the variables used in the analysis

	Mean	SD	%
Student-level variable			
Recommendation	3.1	1.3	
Total score on the school leaver's test	532.5	10.5	
Achievement test (language)	111.5	3.6	
Achievement test (mathematics)	116.7	9.4	
Achievement test (reading comprehension)	54.4	16.5	
Gender			
Boys (reference)			50.1
Girls			49.9
Socio-ethnic background			
High SES			28.3
Middle SES (reference)			35.0
Low-SES Dutch			17.9
Low-SES Turkish or Moroccan			12.7
Low-SES other foreign			6.1
Class-level variable			
Prior achievement (average)	532.4	5.1	
Proportion of students with low-educated parents	0.4	0.3	

Teacher expectations

The teacher's recommendations at the end of primary education were measured using a questionnaire. For each student, the teachers could mark one or more of eight categories that corresponded to the tracks in secondary education (1 = *practical training*; 2 = *basic track with additional support*; 3 = *basic track*; 4 = *middle track*; 5 = *combined track*; 6 = *theoretical track*; 7 = *higher general secondary education*; and 8 = *pre-university education*). The recommendations provided in the questionnaire were converted into scores on the 'educational ladder', a tool developed to map the recommendations and educational positions of students in Dutch secondary education (Bosker & Van der Velden, 1985). This transformation was almost linear, resulting in scores ranging from 1 to 5.

School leaver's test

The students' scores on the school leaver's test (spring 2005) were available in the data set. This high-stakes test was administered by almost 85% of Dutch primary schools. The test consists of three parts: Dutch language (100 items), mathematics (60 items), and information processing (40 items). Students' scores are converted by Cito, a Dutch institute for educational measurement, to a scale ranging from 501 to 550.

Tests from monitoring system

The scores on Dutch language, mathematics, and reading comprehension tests were derived from a monitoring system commonly used by schools to monitor the achievement and progress of their students. The conditions under which these tests were administered by the schools were less stringent than those set for the school leaver's test. The language test consisted of 64 items; the mathematics test, of 120 items; and the reading

comprehension test, of 50 items. The scores used in the analyses were the scores on the underlying latent scales of language, mathematics, and reading comprehension used in the monitoring system.

Demographic characteristics of students

The schools provided the information on the students' backgrounds. A dummy variable was created for gender, where boys formed the reference group. The available socio-ethnic background variable was a nominal variable with five categories based on the level of the parents' education and their ethnicity. Three groups were defined based on parental education, namely low (pre-vocational education), middle (senior secondary education), and high (higher education or university). The students in the lowest category of parental education were split up into three groups based on their ethnicity: (1) Dutch, (2) Turkish and Moroccan, and (3) other foreign students.² Students from the largest group (middle parental education) formed the reference group in the subsequent analysis.

Student populations of classes

Two class composition variables were included: The average score on the school leaver's test and the proportion of students with low-educated parents. These variables were derived by aggregating the students' individual scores to the class level. The aggregated variables were based on all students in a class for whom this information was available. Also the students who were excluded from the analyses on the individual level due to missing values on one or more predictor variables.

Analytic strategy

First, a correlation table was obtained to assess the strength of the associations between teacher expectations and the indices of prior student achievement. Second, the data were analysed using two-level hierarchical models (Snijders & Bosker, 2012) using the MLwiN 2.29 software (Rasbash, Charlton, Browne, Healy, & Cameron, 2009), with students (level 1) nested within teachers (level 2). An unconditional model (Model 0) with teacher expectation as dependent variable was estimated to investigate the size of the differences in teacher expectations among classes. In Model 1, we included the performance variables as predictors of teacher expectations, applying grand-mean centring to the predictor variables (Enders & Tofoghi, 2007). In this way, it was possible to establish whether or not there were teachers whose expectations were generally either too high or too low across all the students in their class. In Model 2, the students' demographic characteristics were included; this provided information on whether the teachers were in general biased in favour of (or against) students with specific demographic characteristics. Next, we

² Following the recommendation of Driessen et al. (2006), only the category of low parental education was split up into further categories because there were considerably fewer students with non-Dutch parents in the middle and high parental education groups. For example, 74% of the students with a Turkish background and 79% of the students with a Moroccan background fall into the category of low-educated parents. Turkish and Moroccan students were selected in a separate category because they represent two large minority groups in the Netherlands that developed since the 1960s. Previous research has indicated that these groups behave rather similarly in the context of education. Students were classified as other foreign if their parents originated from a country other than the Netherlands, Turkey, or Morocco. The category of students from other foreign backgrounds consisted mostly of students whose parents originated from Suriname, the Netherlands Antilles, the former republic of Yugoslavia, China, and Vietnam.

estimated a model in which random slopes were allowed for the demographic variables (Model 3). The aim of this model was to show whether the relationship between the teachers' expectations and the demographic variables differed among the teachers and thus whether teachers differed in specific bias. The random slopes for demographic variables were allowed to correlate among each other and with the general bias in order to estimate a full variance–covariance matrix on the teacher level. In Model 4, we included the two class-level predictors and also the cross-level interactions between these class-level variables and the student-level demographic variables gender and socio-ethnic background, to investigate whether the expectations of the teachers for particular subgroups of students were dependent on the composition of their classes.

Results

Correspondence between teacher expectations and student achievement

In Table 2, the zero-order correlations are presented between teacher expectations and several measures of student performance. A high correlation was found between the expectations of teachers and the scores on the school leaver's test ($r = .89$). Somewhat lower yet significant positive relations were observed between the teachers' expectations and the test scores for language ($r = .65$), mathematics ($r = .73$), and reading ($r = .75$). These high correlations indicate that the expectations of the teachers were closely related to the performance of the students during the final grade of primary education.

General teacher expectation bias

The results of the various multilevel models are presented in Table 3. In Model 0 (unconditional model), 16% (ICC = .16, ICC = Intra Class Correlation) of the variance in teacher expectations is associated with the class level, indicating that there is a substantial clustering of teacher expectations at the class level. The results of Model 1 indicated that both the school leaver's test and the other three achievement tests showed unique positive associations with the teachers' recommendations. The performance variables greatly reduced both the unexplained between-student and the between-teacher variances. The latter indicated that the student populations of the classes in the sample differed with respect to performance. Together, these variables explained 80.2% of the total variance in teacher expectations. Of the remaining variance, 22% (ICC = .22) was associated with the class level, showing that the differences among teachers with respect to their recommendations became more pronounced.

Table 2. Zero-order correlations among teacher expectations and background information on which the teachers could base their expectations

	Teacher expectations	School leaver's test	Language test	Mathematics test
School leaver's test	.89*			
Language test	.65*	.68*		
Mathematics test	.73*	.77*	.51*	
Reading comprehension test	.75*	.79*	.68*	.60*

Note. $n = 7,550$; * $p < .001$.

Table 3. Results from the multilevel regression models of teacher expectations

	Model 0		Model 1		Model 2		Model 3		Model 4	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed Part</i>										
Student-level variables										
Intercept	3.117*	.026	3.112*	.026	3.104*	.016	3.104*	.017	3.101*	.016
School leaver's test			0.081*	.001	0.079*	.001	0.079*	.001	0.080*	.001
Language test			0.017*	.002	0.017*	.002	0.017*	.002	0.017*	.002
Mathematics test			0.018*	.001	0.020*	.001	0.020*	.001	0.020*	.001
Reading comprehension test			0.008*	.001	0.007*	.001	0.007*	.001	0.007*	.001
Low-SES Turkish or Moroccan					-0.038	.022	-0.032	.024	-0.029	.033
Low-SES other foreign					-0.057*	.026	-0.057	.030	-0.062	.033
Low-SES Dutch					-0.113*	.017	-0.111*	.017	-0.104*	.017
High SES					0.082*	.015	0.084*	.015	0.075*	.017
Gender (girls)					0.090*	.012	0.091*	.013	0.091*	.012
Class-level variables										
Class mean school leaver's test									0.022*	.004
Proportion low-education students									-0.265*	.078
Cross-level interaction										
Class mean school leaver's test * Gender									-0.001	.003
Proportion low education * Gender									-0.086	.056
Class mean school leaver's test *									-0.007	.005
Low-SES Turkish or Moroccan									-0.008	.006
Class mean school leaver's test *									-0.004	.004
Low-SES other foreign									-0.004	.004
Class mean school leaver's test *									0.001	.004
Low-SES Dutch										
Class mean school leaver's test *										
High SES										

Continued

Table 3. (Continued)

	Model 0		Model 1		Model 2		Model 3		Model 4	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Proportion low education * Low-SES										
Turkish or Moroccan									-0.104	.119
Proportion low education * Low-SES									-0.074	.136
other foreign										
Proportion low education * Low-SES									-0.114	.091
Dutch										
Proportion low education * High SES									-0.049	.090
<i>Random Part</i>										
Teacher-level intercept variance	0.239	.021	0.064	.005	0.062	.005	0.073	.008	0.052	.004
Student-level intercept variance	1.249	.021	0.230	.004	0.226	.004	0.218	.004	0.225	.004
ICC	0.16		0.22							
<i>Model fit</i>										
-2 * log likelihood	23754.623		11131.557		10968.569		10911.556		10891.084	
Number of classes	500		500		500		500		500	
Number of students	7,550		7,550		7,550		7,550		7,550	

*p < .05.

The between-teacher differences in recommendations from Model 1 are shown in Figure 1. Each triangle and the associated 95% confidence interval in the figure represents a teacher. The value zero in the graph represents the predicted average recommendation of the teachers given the performance of their students. Teachers on the left side of the figure had given their students lower recommendations than expected, while teachers on the right side of the figure had given higher recommendations than might be expected given their students' performance. For most teachers (70.6%), zero was included in the confidence interval, indicating that the observed expectations did not differ significantly from the expected expectations. For the classes where the confidence interval did not include zero, the difference between the observed and the expected teacher expectation was statistically significant (29.4%), indicating general bias (15.2% downward and 14.2% upward). The mean difference between the low-expectation teachers and the high-expectation teachers amounted to one point on the educational ladder, which corresponds to a difference of one track in Dutch secondary education.

Specific teacher expectation bias

The results of the multilevel model, in which the demographic variables gender and socio-ethnic background were included, are presented in Model 2 (Table 3). Compared with Model 1, the coefficients of the performance variables have hardly changed; however, there were significant unique associations between teacher expectations and both gender and socio-ethnic background. When achievement was controlled for, girls received higher track recommendations from their teachers than boys ($b = .90$). Furthermore, students with higher educated parents received higher track recommendations than students with middle-educated parents ($b = .082$). No significant differences were found between the reference group and Turkish and Moroccan students from low-SES families. Dutch and other foreign students from low-SES families were given lower track recommendations than the reference group ($b = -.113$ and $b = -.057$, respectively).

In Model 3, we added random slopes for the demographic variables to the model (see Table 3 for fixed part; Table 4 for random part). Including random slopes in the model for the demographic variables (gender and socio-ethnic background) led to a significant

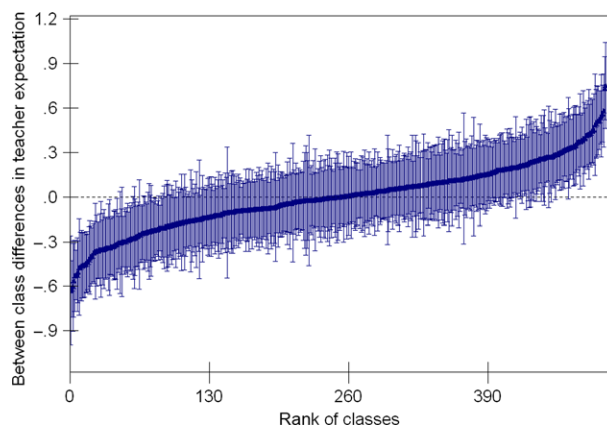


Figure 1. Differences among teachers in level of expectations after differences in student performance are controlled for.

increase in model fit ($\chi^2 = 57.01$; $df = 21$; $p < .005$). In Table 4, the variances surrounding the slopes are presented on the diagonal (light grey-coloured cells). The slope variances, indicating differences in specific bias among teachers for the demographic groups of students, are also presented in Figure 2. Although the teachers' recommendations for girls generally tended to be somewhat higher than those for boys, this difference was not similar across all teachers ($\sigma^2 = .015$). The observed range of differences across teachers between boys' and girls' recommendations lay between $-.08$ and $.35$, a negative coefficient indicating bias in favour of boys. Furthermore, the differences among teachers in specific bias against Dutch students with low-educated parents were relatively small ($\sigma^2 = .007$). The observed range of differences in the recommendations between the reference group and Dutch students with low-educated parents lay between $-.21$ and $-.02$. This implies that these students got lower recommendations than the reference group in all classes. For students from Turkish and Moroccan backgrounds with low-SES families, the observed bias varied between $-.29$ and $.19$ ($\sigma^2 = .024$). This indicates substantial differences between teachers in their bias, with some teachers in favour of students from Turkish and Moroccan backgrounds with low-SES families and other teachers in favour of the reference group. Similarly, the differences among teachers in respect of other foreign students from low-SES families ($\sigma^2 = .035$; range from $-.47$ to $.22$) were relatively large. For the two latter groups, the differences among teachers were greater than the average bias taking all teachers together.

We also found some interesting patterns of associations between the specific expectation biases for several subgroups of students. The covariances (below diagonal) and correlations (above diagonal) between the teacher-level residuals are presented in Table 4. The first pattern relates to the association of bias in expectations for the middle- and low-SES student subgroups. Teachers who are relatively positively biased towards the reference group tend to be biased in favour of students from the low-SES subgroups. However, there seemed to be some differences in the strength of these correlations (low-SES Dutch: $r = .297$; low-SES Turkish or Moroccan: $r = .026$; and low-SES other

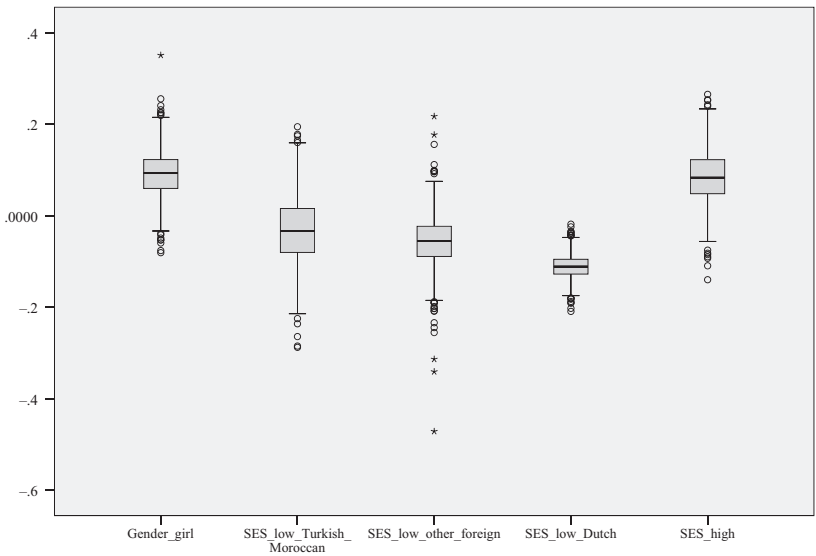


Figure 2. Differences among teachers for demographic groups of students based on gender and socio-ethnic background.

Table 4. Variance–covariance matrix and correlations for the random part of Model 3 on the teacher level (random slopes multilevel model)

Variance–covariance matrix	Intercept	Low-SES Dutch	Low-SES Turkish or Moroccan	Low-SES other foreign	High SES	Gender (girls)
Intercept	.073 (.008)	.297***	.026	.460***	–.874***	–.349***
Low-SES Dutch	–.010 (.005)	.007 (.008)	.532***	.471***	–.250***	–.657***
Low-SES Turkish or Moroccan	.001 (.009)	.009 (.010)	.024 (.014)	–.10	–.125**	–.245***
Low-SES other foreign	.003 (.011)	.010 (.011)	.001 (.013)	.035 (.019)	–.297***	.091*
High SES	–.014 (.006)	.002 (.005)	.009 (.010)	.002 (.010)	.001 (.005)	.169***
Gender (girls)	–.010 (.005)	–.005 (.004)	–.004 (.006)	.003 (.008)	–.001 (.004)	.015 (.005)

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

foreign: $r = .460$). Similarly, teachers who are biased in favour of the Dutch low-SES students tended to be biased in favour of the other low-SES student subgroups as well (Dutch and Turkish/Moroccan: $r = .532$; Dutch and other foreign: $r = .471$). However, the association between bias for the Turkish/Moroccan and bias for other foreign students was very small and non-significant ($r = -.010$). This implies that teachers who were biased against Turkish/Moroccan students were not necessarily biased towards other foreign students.

A second apparent pattern relates to the association of bias in expectations for the high-SES subgroup and the other SES subgroups. Teachers who were biased in favour of the high-SES subgroup tended to be biased against students from the middle- and low-SES subgroups (middle SES: $r = -.874$; low-SES Dutch: $r = -.250$; low-SES Turkish or Moroccan: $r = -.125$; and low-SES other foreign: $r = -.297$). In general, these negative correlations imply that some teachers exaggerated the differences among subgroups (relatively positively biased expectations for the high-SES group), while other teachers downplayed the differences among the SES groups (relatively negatively biased expectations for the high-SES group). Although all these correlations are negative, their strength differed. The association that teachers who were biased towards the high-SES subgroup tended to be biased against the reference group (middle SES) is by far the strongest. The association between the teacher-level residuals of the middle-SES and high-SES subgroups is presented in Figure 3. The residuals of half of the teachers ($n = 250$) can be found in the upper left quadrant, indicating bias in favour of the high-SES subgroup and bias against the middle-SES subgroup. For these teachers, the differences between their expectations for the high-SES students and the low-SES students were larger than expected based on the students' performance (exaggerated). Another large number of residuals ($n = 208$) can be found in the upper right quadrant of the figure, which indicates bias in favour of both groups of students. Finally, the residuals of 42 teachers can be found in the lower right quadrant, indicating bias in favour of the reference group and bias

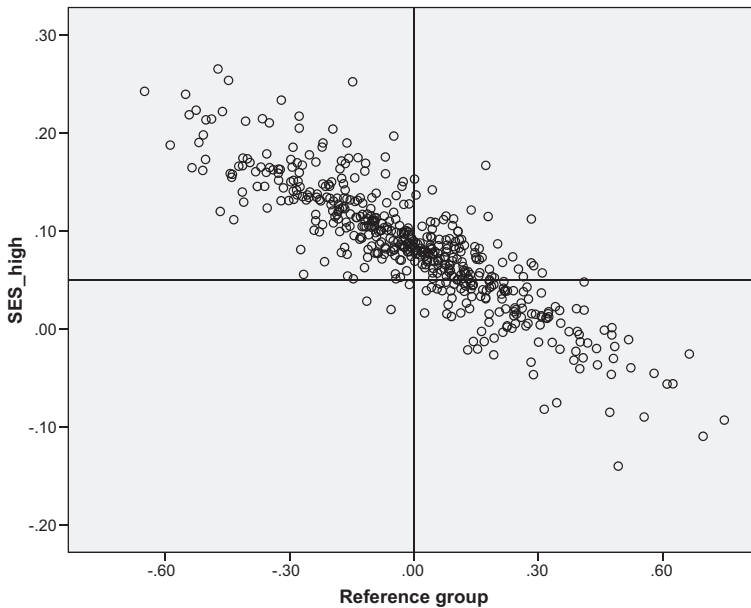


Figure 3. Association between the teacher-level residuals for the reference group (middle-SES) and the high-SES subgroup.

against the high-SES subgroup. For these teachers, the difference in teacher expectations for the middle-SES and high-SES groups of students was smaller than the students' performance predicted (downplayed).

Finally, there is no clear pattern among the correlations between gender bias and SES subgroup bias.

Explaining the differences in teacher expectations in relation to classroom context

Model 4 presents the results of including the two class-level variables as well as the cross-level interactions in the multilevel regression model (Table 3). The fit of Model 4 appeared to be better than that of Model 2 ($\chi^2 = 77.49$; $df = 12$; $p < .001$). When all student characteristics were taken into account, the students tended to receive higher recommendations if they were in a high-performing class and/or in a class with relatively few students from low-SES families. We also found that the fixed coefficients indicating bias in favour of or against the different socio-ethnic groups decreased due to the inclusion of the class-level variables. None of the ten cross-level interactions between the class-level variables and the demographic student-level characteristics appeared to be significant. This indicates that the differences between the reference groups and the other demographic groups were relatively independent of the class population.

Discussion

In the current study, we investigated (1) whether teachers generally were less accurate in their expectations for subgroups of students based on their demographic characteristics and after the performance of the students is taken into account, (2) whether there were

between-teacher differences in their expectations for students with similar performance records, and (3) whether the teachers' expectations were associated with the class composition. Differences between teachers, both in the average level of their expectations and their specific expectations for demographic subgroups, were investigated in a sample of 7,550 students in 500 classes in the final grade of Dutch primary education.

Our first finding is that prior achievement variables explained 80% of the variance in teacher expectations at the student level. This indicates a close correspondence between the teachers' academic expectations for the students and the students' actual performance, which is consistent with our first hypothesis (H1) and the conclusion of several previous researchers (Brophy, 1983; Jussim & Eccles, 1992; Jussim & Harber, 2005; Luyten & Bosker, 2004). The expectations of the Dutch teachers can be considered fairly accurate. At the teacher level, we found that for more than 70% of the teachers in the sample, the average observed expectation did not differ significantly from the average expected expectation based on the performance records of the students in their classes. Given that the teachers in this sample had access to students' performance information before formulating their expectations, that they were supposed to formulate a well-founded expectation, and that the expectations were given at the end of the school year, high correlations among expectations and performance were expected (Ready & Wright, 2011).

Consequently, in the subsequent multilevel analyses, 21.8% of the total variance in expectations was associated with the teacher level, after student performance was controlled for. These results were in line with our hypothesis (H3), namely that teachers differ in the extent of their general bias (i.e., systematically expecting too much or too little of all students in their class). The between-teacher differences were somewhat smaller in the current study than in the previous research in the Netherlands (De Boer *et al.*, 2010). This divergence may originate from several methodological differences between the two studies, such as the sampling (secondary schools vs. primary schools) and the variables taken into account (school leaver's test, intelligence and performance motivation vs. school leaver's test and tests from monitoring systems). However, the findings of both studies suggest that there are considerable differences between teachers in their expectations for students with equal performance records.

A second finding was that for the Dutch context, some indications of teacher expectation bias based on gender and socio-ethnic background were found, after student performance was controlled for. This finding is similar to those of previous research (Glock & Krolak-Schwerdt, 2013; McKown & Weinstein, 2008; Ready & Wright, 2011; Rubie-Davies *et al.*, 2006; Speybroeck *et al.*, 2012). Among students with equal performance records, higher teacher expectations were observed for girls and for students from more affluent families. This latter pattern confirmed our second hypothesis that teachers are specifically biased against students from low socio-economic backgrounds. These results may indicate that the general differences in teacher expectations for students from more and less affluent families (i.e., taking all teachers together) may add to the already existing achievement gap between these groups, with unfavourable consequences for the subsequent school careers of those in the latter group. This is especially important in a context where teacher expectations are an important criterion for the students' placements in secondary education tracks, as is the case in the Netherlands (Driessen, 2005). Furthermore, we found the lowest expectations for Dutch students, which is similar to the previous research findings in the Netherlands (Claassen &

Mulders, 2003). Given the results of previous studies, we did not expect to find differences among ethnic groups (H2). It is therefore particularly interesting that the minority students from low socio-economic backgrounds received higher recommendations compared to their Dutch peers with similar low socio-economic backgrounds.

The third finding of this study is that we found that teachers differed considerably in the extent of their bias towards or against specific demographic groups, thereby confirming the fourth very tentative hypothesis. A striking result is that the differences between teachers in specific bias towards Turkish and Moroccan and other foreign students from low-SES families were larger than the average bias in the sample (i.e., taking all teachers together). This finding implies that in general there is almost no bias towards these groups of students, but that the accuracy of the expectations for these students is dependent on the individual teacher, as some teachers are positively biased while others are negatively biased. For Dutch students from low-SES families, the average bias in the sample was larger than the between-teacher differences. In general, lower recommendations were found for these students, and they were found in all classes. For this particular group, teachers appeared to be relatively consistently negatively biased, indicating that the expectations were less dependent on the individual teacher or the composition of the class. There may be several explanations for the latter result. First, Dutch students from low-SES families are mostly of the same ethnic background as the reference group (which includes only a relatively small proportion of non-Dutch students), but differ in one key demographic feature. As a consequence, comparisons between the reference group and Dutch students from low-SES families can be made relatively easily. In contrast, any differential performance by non-ethnic Dutch students might (perhaps unconsciously) be attributed to other (social cultural) factors (which may be positive or negative). Second, the ongoing political debate in the Netherlands may have created a pressure to behave politically correctly, which may have influenced the teachers' recommendations for the Turkish, Moroccan, and other foreign students. The accuracy of recommendations for ethnic minority students has been the focus of this debate, and little attention has been paid to the recommendations given to Dutch disadvantaged students (Claassen & Mulders, 2003). This may explain the larger heterogeneity in teacher expectations for ethnic minority groups.

The fourth important finding relates to a pattern that was apparent on the teacher level, showing that teachers who were biased in favour of one of the low-SES student subgroups tended to be biased in favour of the other low-SES and the middle-SES student subgroups as well. However, teachers who were biased in favour of the high-SES subgroup tended to be biased against students from the middle- and low-SES subgroups. The latter finding is of particular importance in the context of teacher expectations possibly contributing to the already existing achievement gap between high- and low-SES subgroups. The possibility of teacher expectations contributing to the achievement gap is more plausible for students whose teachers exaggerated the differences among subgroups, while it is less plausible for students whose teachers played down the differences among SES groups.

Finally, a fifth finding is that the recommendations teachers gave to students were dependent not only on the individual characteristics of the students but also on the characteristics of the other students in the same class. Teachers tended to have higher expectations of children in high-achieving classes and in classes with relatively few children from low-SES families, after controlling for differences in prior achievement and individual student background. This finding seems to correspond with those of

previous research (McKown & Weinstein, 2008; Ready & Wright, 2011). However, our findings also revealed that teacher expectation bias towards or against specific subgroups was independent of the class population. This result contrasts with our hypothesis (H4) and the findings of McKown and Weinstein (2008), who concluded that bias against specific ethnic groups was greater in more heterogeneous classrooms. Differences between the results of this study and the study by McKown and Weinstein may arise, among other things, from differences in the operationalization of the class composition variables. Our operationalization (average performance and proportion of students from low-SES backgrounds) did not directly relate to perceivable contrasts between groups of students within classes, whereas McKown and Weinstein (2008) used the number of ethnic groups which directly relates to manifest contrasts within classes.

In interpreting the results of this study, a number of limitations need to be considered. First, the study was conducted in the particular context of Dutch primary education and used information that was gathered at the end of the school year close to a transition to a highly tracked secondary education system. Several findings, such as the degree of association among performance and the teachers' expectations of students and the extent of specific teacher expectation bias, may depend on several characteristics of the sample and context, such as the educational system, heterogeneity of classes, and the prevalence and stereotyping of ethnic minority groups. Furthermore, the results derived from the estimated models can only be interpreted in the context of the available performance information, namely the school leaver's test and achievement tests from the schools' monitoring systems. It remains questionable whether these two sources, although reliable, were sufficient to calibrate the teacher expectations and whether other valid sources of information that teachers may use in forming expectations were omitted. In addition, we applied a commonly used method to investigate accuracy (Cooper *et al.*, 1982; De Boer *et al.*, 2010; Hinnant *et al.*, 2009), by relating teacher expectations to the students' previous performance records. Furthermore, in order to investigate whether teacher expectations predict, but do not cause, students' future academic achievement (Jussim, 1991), a common definition of expectation accuracy, at least two measurements of performance need to be assessed. In the current study, we could only investigate whether the teachers' expectations were accurate reflections of previous performance. Therefore, we cannot make strong statements about the predictive validity of the expectations, as a follow-up measurement of performance was not available.

All in all, our findings imply that only investigating average trends while ignoring differences between teachers does not do justice to the complex nature of teacher expectations; the recommendations (expectations) students receive depend on many factors, such as the students' own performance, their gender and socio-ethnic background, the performance and socio-ethnic composition of their class, and also on their individual teacher. Given the large differences that were observed among teachers in their expectation biases (general and specific), obtaining more (or better) answers as to why teachers' expectations differ and under what circumstances teachers develop biased expectations (McKown & Weinstein, 2008) is necessary. Furthermore, research into the differential classroom behaviour of teachers corresponding with having high or low general expectations (Rubie-Davies, 2007, 2010) may be of particular importance. Multilevel models should therefore be used, not only because of the nested structure of educational data, but also to analyse in more detail the complex nature of between-teacher differences.

References

- Bosker, R. J., & Van der Velden, R. K. W. (1985). *Schooleffecten en rendementen [School effectiveness and efficiency]*. Groningen, the Netherlands: RION: Instituut voor Onderwijsonderzoek Rijksuniversiteit Groningen.
- Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75*, 631–661. doi:10.1037/0022-0663.75.5.631
- Claassen, A., & Mulders, L. (2003). *Leerlingen na de overstap. Een vergelijking van vier cohorten leerlingen na de overgang van basisonderwijs naar voortgezet onderwijs met nadruk op de positie van doelgroep leerlingen van het onderwijsachterstanden beleid [Students after the transition. A comparison of four student cohorts after the transition from primary to secondary education with special focus on the position of students from minority groups and low ses families]*. Nijmegen, the Netherlands: ITS, Radboud Universiteit Nijmegen.
- Cooper, H., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology, 74*, 577. doi:10.1037/0022-0663.74.4.577
- De Boer, H., Bosker, R., & Van der Werf, M. P. C. (2007). De gevolgen van onder- en overadvisering. In Inspectie van het Onderwijs (Ed.), *Onderadvisering in beeld [Low track recommendations in the picture]* (pp. 83–92). Utrecht, the Netherlands: Inspectie van het Onderwijs.
- De Boer, H., Bosker, R. J., & Van der Werf, M. P. C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology, 102*, 168–179. doi:10.1037/a0017289
- Driessen, G. (2005). De totstandkoming van de adviezen voortgezet onderwijs: Invloeden van thuis en school [The realisation of track recommendations: Home and school influences]. *Pedagogiek, 25*, 279–298.
- Driessen, G. (2011). *Onderadvisering van allochtone leerlingen? [Low track recommendations for minority students?]*. Nijmegen, the Netherlands: ITS, Radboud Universiteit Nijmegen.
- Driessen, G., & Smeets, E. (2007). De relatie tussen prestaties en advies: Onder- of overadvisering bij de overgang van basis- naar voortgezet onderwijs? [The association between student performance and track recommendations. Low or high recommendations for the transition from primary to secondary education?] In Inspectie van het Onderwijs (Ed.), *Onderadvisering in beeld* (pp. 59–81). Utrecht, the Netherlands: Inspectie van het Onderwijs.
- Driessen, G., van Langen, A., & Vierke, H. (2006). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Zesde meting 2004/05 [Primary education: Data collection report for student data and parent questionnaires. Report PRIMA cohort study, sixth wave 2004/2005]*. Nijmegen, the Netherlands: ITS.
- Dronkers, J., van Erp, M., Robijns, M., & Roeleveld, J. (1998). Krijgen leerlingen in de grote steden en met name in Amsterdam te hoge adviezen? De relaties tussen taal- en rekenscores en advies binnen en buiten de Randstad onderzocht [Do students from Amsterdam and other big cities in the Netherlands receive high track recommendations? An investigation of the association between language and mathematics scores and recommendations in and outside the Dutch largest cities]. *Tijdschrift voor Onderwijsresearch, 17*–30.
- Dusek, J. B., & Joseph, G. (1982). Bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology, 75*, 327–346. doi:10.1037/0022-0663.75.3.327
- Enders, C. K., & Tofoghi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121–138. doi:10.1037/1082-989X.12.2.121
- Glock, S., & Krolak-Schwerdt, S. (2013). Does nationality matter? The impact of stereotypical expectations of 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

- Helwig, R., Anderson, L., & Tindal, G. (2001). Influence of elementary student gender on teachers' perceptions of mathematics achievement. *The Journal of Educational Research*, 95, 93–102. doi:10.1080/00220670109596577
- Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology*, 101, 662–670. doi:10.1037/a0014306
- Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of the literature. *Review of Educational Research*, 59, 297–313. doi:10.3102/00346543059003297
- Inspectorate of Education (2007). *Teacher recommendations in the picture. [Onderadvisering in beeld]*. Utrecht, the Netherlands: Author.
- Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review*, 98, 54–73.
- Jussim, L., & Eccles, J. S. (1992). Teacher expectations II: Construction and Reflection of students achievement. *Journal of Personality and Social Psychology*, 63, 947–961. doi:10.1037/0022-3514.63.6.947
- Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 281–388). Bingley, UK: Emerald Group.
- 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
- Luyten, H., & Bosker, R. J. (2004). Hoe meritocratisch zijn schooladviezen? [to what extent are track recommendations meritocratic?] *Pedagogische Studiën*, 81(2), 89–103.
- Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, 72, 791–809. doi:10.1037/0022-3514.72.4.791
- Madon, S., Willard, J., Gyll, M., & Scherr, K. C. (2011). Self-fulfilling prophecies: Mechanisms, power and links to social problems. *Social and Personality Psychology Compass*, 5, 578–590. doi:10.1111/j.1751-9004.2001.00375.x
- McKown, C., & Weinstein, R. S. (2002). Modeling the role of child ethnicity and gender in children's differential response to teacher expectations. *Journal of Applied Social Psychology*, 32, 159–184. doi:10.1111/j.1559-1816.2002.tb01425.x
- McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, 46, 235–261. doi:10.1016/j.jsp.2007.05.001
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8, 193–210.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2009). *MLwiN Version 2.1*. Bristol: Centre for Multilevel Modelling, University of Bristol.
- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities The role of child background and classroom context. *American Educational Research Journal*, 48, 335–360. doi:10.3102/0002831210374874
- Riley, T., & Ungerleider, C. (2012). Self-fulfilling prophecy: How teachers' attributions, expectations, and stereotypes influence the learning opportunities afforded Aboriginal students. *Canadian Journal of Education*, 35, 303–333.
- Roeleveld, J., Driessen, G., Ledoux, G., Cuppen, J., & Meijer, J. (2011). *Doelgroepopleidingen in het basisonderwijs. Historische ontwikkeling en actuele situatie [At risk students in Dutch primary education. Historical developments and the current situation]*. Amsterdam, the Netherlands: Kohnstamm Instituut.
- Rosenthal, R., & Jacobsen, L. (1968). *Pygmalion in the classroom: Teacher expectations and pupil's intellectual development*. New York, NY: Holt, Rinehart & Winston.
- Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high and low expectation teachers. *British Journal of Educational Psychology*, 77, 289–306. doi:10.1348/000709906X101601

- Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology*, *80*, 120–135. doi:10.1348/000709909X466334
- Rubie-Davies, C. M., 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
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London, UK/Thousand Oaks, CA/New Delhi, India: Sage.
- Social and Cultural Planning Office (2005). *Jaarrapport integratie 2005 [Annual report integration 2005]*. The Hague, the Netherlands: Author.
- Sorhagen, N. S. (2013). Early teacher expectations disproportionately affects poor children's high school performance. *Journal of Educational Psychology*, *105*, 465–477. doi:10.1037/a0031754
- Speybroeck, S., Kuppens, S., Van Damme, J., Van Petegem, P., Lamote, C., Boonen, T., & De Bilde, J. (2012). The role of teachers' expectations in the association between children's SES and performance in Kindergarten: A moderated mediation analysis. *PLoS One*, *7*, e34502. doi:10.1371/journal.pone.0034502
- 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*, 689–705. doi:10.1037/0022-0663.99.2.253
- Timmermans, A. C., Kuyper, H., & van der Werf, M. P. C. (2013). *Schooladviezen en onderwijsloopbanen: Voorkomen, risicofactoren en gevolgen van onder- en overadviesing [Track recommendations and educational careers; Prevalence, risk-factors and consequences of high and low track recommendations]*. Groningen, the Netherlands: GION.
- Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, *47*, 497–527. doi:10.3102/0002831209353594