



Always look on the bright side of students: does valence of teacher perceptions relate to students' educational performance?

Nancy Tandler¹ · Claudia Dalbert²

Received: 4 June 2019 / Accepted: 10 June 2020 / Published online: 25 June 2020
© The Author(s) 2020

Abstract

We studied the role of the valence of female teachers' perceptions in predicting various student outcomes. Using a prospective design, at the beginning of Grade 5 (T1), we assessed evaluations by 43 female teachers of either mathematics or German, belonging to 22 secondary schools in Germany. We combined a qualitative measurement by using an open-ended question, with a quantitative data analysis. At T1 and T2 (4 months later), the assigned students ($N=635$) reported their subject-specific achievement, learning and performance goals, and quality of relationship to their teacher. Multilevel analyses revealed that positive teacher perceptions increased future student achievement, while negative teacher perceptions reduced immediate and future learning and performance goals. Further, positive teacher perceptions explained differences in student outcomes between classes, whereas negative teacher perceptions explained differences in student outcomes between schools.

Keywords Educational attainment · Longitudinal design · Multilevel analyses · Self-fulfilling prophecies · Teacher expectations · Teacher perceptions

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11218-020-09573-z>) contains supplementary material, which is available to authorized users.

✉ Nancy Tandler
nancy.tandler@psych.uni-halle.de

¹ Department of Psychology, Martin-Luther University Halle-Wittenberg, 06099 Halle (Saale), Germany

² Department of Educational Psychology, Martin-Luther University Halle-Wittenberg, 06099 Halle (Saale), Germany

1 Introduction

Teachers' perceptions (TPs) of student attributes are critical for students' learning experience and educational success. TPs appear to be related to how teachers adapt their classroom behavior to students' needs (Rubie-Davies 2006, 2007), and they influence teachers' instructional decisions and placement of students in particular ability learning groups (Ready and Wright 2011). Teachers' perceptions of particular student outcomes, such as achievement, have been studied extensively (e.g., Rubie-Davies 2007), whereas effects on students' social and motivational outcomes have been considered less frequently. In order to narrow this gap in the literature, we tested the associations between teachers' general perceptions of students and three student outcomes: their achievement, their goal orientations, and the quality of relationships with their teachers. In previous research on teacher perceptions, teachers were predominantly requested to rate particular student outcomes. We want to extend on this by examining whether teachers' general ways of perceiving and approaching students in daily life are also relevant for student outcomes, when teachers are not explicitly asked to rate students' attributes. In particular, we were interested in the associations involving teachers' general perceptions of students as a teacher-centered variable (e.g., Rubie-Davies 2006, 2007). We operationalized teacher perceptions as teachers' social cognitions when freely describing students according to their own understanding, whereas we distinguish teacher perceptions from teacher expectations in the sense that teacher perceptions can be understood as antecedents to teacher expectations (Hofer 1986) or in line with Brophy and Good (1974) as a more generalized aspect of teacher expectations.

When teachers freely describe students, many aspects can be examined, such as the content, coherence, stability, and valence of teacher cognitions. Valence is defined as the positive or negative psychological value that one person assigns to another (Wirtz and Strohmer 2014) and constitutes the focus of the current study. The valence of teachers' cognitions about students seems to be an important aspect of teacher judgments (e.g., Givvin et al. 2001). Research has shown that students can recognize teachers' attitudes towards them (e.g., Rubie-Davies 2006) and that these attributes contribute to student learning outcomes (Wentzel 1999). The aim of the present study was to examine the valence of TPs and to investigate whether this is associated with different student outcomes. Previous studies have shown that student outcomes can vary across classes, schools, or both. Therefore the present study used a multilevel approach including class and school levels, to examine whether the associations of valence of TPs with students' educational attainment varied systematically between classes and/or schools.

1.1 Previous research on teachers' social cognitions

Teachers' perceptions of students can be described as social cognitions. Social cognitions explain how people process social information, in particular, how people select, store, and use information about others (Hamilton 2005). During this

mental process, people use various cognitive strategies that help them to process new social information and to fill in social information where it is missing. People can apply mental structures (e.g., implicit personality theories) that help them to categorize social information (Hofer 1986), and these mental structures can be related to more specific social cognitions such as expectations, opinions, and decisions about others, which may influence how those others are treated (Brophy and Good 1974; Hofer 1986).

In schools, a prominent example of teachers' social cognitions about students are teacher expectations. Teacher expectations can be defined as teachers' conscious or unconscious perceptions of students' future ability to succeed (Dusek and Joseph 1983). Initially, Rosenthal and Jacobson (1968) demonstrated that regardless of students' actual cognitive abilities, high teacher expectations could lead to an increase in students' performance.

Overall, teacher expectations are fairly accurate, though this accuracy seems to vary across studies. Jussim (2012) summarized several studies (e.g., Jussim and Eccles 1992; Kuklinski and Weinstein 2001; Trouilloud et al. 2002) and found that 65–100% of the correlation between teacher expectations of students' achievement and students' actual achievement reflected real differences in ability between students, whereas the remaining 0–35% of the correlation reflected self-fulfilling prophecies, that is, the fact that teacher expectations influenced, rather than predicted, the students' outcomes.

Self-fulfilling prophecies (Merton 1948; Rosenthal and Jacobson 1968) can be defined as predictions that, when they become known, themselves bring about the predicted event. This mechanism probably accounts for some of the predictive power of teacher expectations for student achievement. Because teachers' expectations can overestimate or underestimate students' potential achievement, they may increase or reduce students' actual achievement.

Some studies have investigated the conditions under which self-fulfilling prophecies can occur. Traditionally, such studies have investigated differences in teacher expectation effects according to characteristics or attributes of individual students such as performance level, social background, or minority status (e.g., Jussim et al. 1996; Madon et al. 1997). Whereas, the effects of teacher expectations on student achievement were probably mediated by teachers creating a warmer socio-emotional climate, providing more input, giving more detailed feedback and interacting more with students for whom they had higher performance expectations (Harris and Rosenthal 1985).

However, teacher behaviors addressed to the whole class (e.g., creating a warm socio-emotional climate) rather than to individual students (e.g., praising high-expectation students) appeared to mediate teacher expectations to a greater extent (Harris and Rosenthal 1985). This supports Brophy's (1985) assertion that class-level expectations have more impact on outcomes than student-level expectations. He therefore proposed studying characteristics of teachers rather than characteristics of students. Taking into account characteristics of the perceiver is consistent with the approach of interpersonal perception (Funder's Realistic Accuracy Model 2012).

1.2 Teacher-centeredness of teachers' social cognitions

Studies examining teacher expectations at the class level have concluded that expectations are teacher-centered, that is, some teachers have high expectations whereas others have low expectations for all students in their classroom (e.g., Brattesani et al. 1984; Rubie-Davies 2007). Rubie-Davies (2006) identified high and low expectation teachers. These teachers either had on average high or low expectations for students' future achievement in their classroom relative to students' current performance (low, medium, high). Improvement in students' reading achievement between the beginning and end of a school year were very large ($d=1.01$) for classes of high expectation teachers, but low ($d=0.05$) for classes of low expectation teachers (Rubie-Davies 2010). Teacher practices of high expectation teachers were described as: including all students in challenging and interesting activities, creating a more pleasant classroom climate and warmer student–teacher relationships, managing student behavior more positively, setting mastery goals with each student, providing clear and frequent feedback about student progress, and promoting student autonomy and motivation (Rubie-Davies 2007).

Teacher expectations have also been studied at the school level. These studies considered teachability expectations, defined as school-wide teacher beliefs about students' school-adjustment and cognitive-motivational behaviors. Lower teachability expectations were associated with higher rates of students' self-reported school misconduct (e.g., being late, cheating on tests, doing drugs in school; Demanet and Van Houtte 2012) and indirectly with higher levels of students' math achievement via students' sense of academic futility (Agirdag et al. 2013).

1.3 Teachers' perceptions of students

Teacher perceptions that are not explicit teacher judgments about specific student outcomes but rather reflect teachers' general social cognitions about students may also be predictive of students' educational success.

Previous research on teacher judgments has shown that they are often very different from students' own judgments. Givvin et al. (2001) measured perceived ability, learning goals, positive emotions and negative emotions in mathematics. They found that at the beginning of the year teacher judgments were similar to student self-reports for only two of these four scales (perceived ability, learning goals), and even then the association was weak. There were also qualitative differences between teacher and student judgments. Teacher judgments were more global, that is, they tended to be the same for all four scales, whereas students gave themselves more varied ratings across the four scales. Teacher judgments were also more stable over time, whereas student self-reports changed more over time. Moreover, it was concluded that teachers "tended to see students as generally positive or negative" (Givvin et al. 2001, p. 329). Further, in mathematics and reading comprehension, teachers' global judgments of student performance were more predictive of specific performance measures than were teachers' actual judgments of those specific performance measures (Rausch et al. 2016). Therefore, it is interesting to estimate the

predictive power of teachers' perceptions of students in general, and in particular of the valence of those perceptions.

Teachers' perceptions of students in general seem to be especially relevant when teachers are not instructed to rate particular student outcomes but instead freely describe students according to their own understanding (elicited using open-ended questions; Hofer 1986). These may best reflect teachers' general way of perceiving and approaching students in daily life. There are many classroom situations in which teachers are not explicitly asked to judge their students, yet the valence of teachers' perceptions might still matter. Indeed, it has already been shown that students can recognize teachers' attitudes (e.g., Rubie-Davies 2006), and that these attitudes contribute to student learning outcomes (Ashton 1985; Wentzel 1999). Thus, the valence of teachers' social cognitions seems to matter. For instance, Rubie-Davies (2010) found that high expectation teachers' positive perceptions of students' attributes (e.g., reaction to new work, self-esteem, home environment, relationship with peers) were consistently positively associated with students' outcomes. In contrast, low expectation teachers held both negative and positive attitudes towards students, and these attitudes were only very weakly related to student outcomes. This might indicate that students can distinguish the valence of teachers' social cognitions. It is therefore relevant to examine favorable and unfavorable teacher perceptions.

Assessing teachers' social cognitions using an open-ended question method (Hofer 1986) facilitates recording their positive and negative valences simultaneously and helps to do justice to possible differences in teachers' expressiveness in the evaluation of different student attributes. Previous studies have tended to assess teacher attitudes with closed questions, leading to a neglect of expressiveness (Rubie-Davies 2010). However, there is evidence that teachers express their expectations about students in many different ways and that high expectation teachers express their positive expectations more than low expectation teachers express their negative expectations in classroom behaviors (e.g., Rubie-Davies 2007). For instance, high expectation teachers provided students with more praise for a correct answer and utilized more positive behavior management statements than low expectation teachers, but low expectation teachers were not found to apply more criticism or negative behavior management statements.

1.4 Students' educational attainment

Studies of teacher expectations have primarily focused on student achievement as the outcome of interest (e.g., Rubie-Davies 2007), although a few have also attempted to measure other outcomes (e.g., misconduct in school; Demanet and Van Houtte 2012). It is therefore important to investigate the predictive power of teacher cognitions on other outcomes.

Students' goal orientations can be classified as either learning or performance orientations, and either approach or avoidance orientations (Elliot 1999; Elliot and Church 1997; Elliot and McGregor 2001). Learning-approach goals involve striving to extend one's current knowledge, and learning-avoidance goals involve striving not to lose one's competencies. Performance-approach goals involve trying to

outperform others and to demonstrate competence, whereas performance-avoidance goals involve trying to avoid failure and appearing incompetent. Learning goals (mostly learning-approach goals) seem to be most adaptive for students. Performance-approach goals have been found to be less adaptive, and performance-avoidance goals have been consistently found to be associated with poor outcomes (e.g., Dweck and Leggett 1988; Elliot and McGregor 2001; Finney et al. 2004). In planning our study, we selected student outcomes for which we could derive tentative hypotheses based on previous research. Thus, we selected the dimensions best known for positive student outcomes, that is learning-approach goal orientation, and those best known for having the most negative effects, that is performance-avoidance goal orientation (e.g., Elliot and McGregor 2001). Teachers with high expectations of students were found to provide students with feedback and challenging learning opportunities (Rubie-Davies 2007), which in turn was found to be positively associated with students' learning goal orientations (Lazarides et al. 2018). Low expectation teachers were found to make more procedural statements than teaching statements (Rubie-Davies 2007), which in turn might foster performance goals.

The student–teacher relationship is a crucial outcome given the central role of teachers in the educational system (e.g., Anderman and Anderman 1999; Wentzel 1999). Students' evaluation of their relationship with their teachers is positively related to current and future adjustment to school, as measured by students' personal task goals, feelings of school belonging, positive school affect, achievement (e.g., Roeser et al. 1996), and academic self-efficacy (e.g., Agirdag et al. 2012; Roeser et al. 1996). High expectation teachers were found to create socio-emotional environments that contained more positive behavior management statements and more constructive feedback than low expectation teachers (Rubie-Davies 2007).

2 The present study

An appropriate way of testing group-level associations between teacher cognitions and students' educational attainment is to use multilevel modeling. First of all, multilevel modeling allows testing for class effects, that is, variation of student outcomes across classes. Previous research has already demonstrated that student performance varies across classes (Rubie-Davies 2006) in line with teachers' expectations. However, for student performance, student motivation (Wang and Eccles 2013) or school misconduct, (Demanet and Van Houtte 2012) school effects have also been found, that is, these outcomes varied across schools in line with students' perceptions of the school environment and teachers' expectations about students' teachability. This raises the question of what factors determine these class and/or school variations. Every class or school provides a context that makes students more similar to each other than to students from different contexts. Presumably, there are specific characteristics that contribute to these context effects. For example, teacher cognitions may function as typical classroom characteristics, having similar effects on students attending the same class (e.g., Butler 2012; Rubie-Davies 2006).

Furthermore, teacher cognitions can also serve as school characteristics, making students from one school more similar to other students from that school, because school

contexts can lead to teachers at the same school having similar cognitions of students (e.g., Agirdag et al. 2013; Demanet and Van Houtte 2012; Eyal and Roth 2011). In sum, previous results indicate that student outcomes can vary across different levels, that is, across classes, schools, or both. Therefore, the present study included class and school levels, to test whether the influence of teachers' cognitions on students' educational attainment varied systematically between classes and/or schools, using multilevel modeling.

Instead of gathering data on teacher perceptions for every single student in their class, we took a random sample of each teacher's general tendency to perceive students positively or negatively by asking each teacher to describe three students. We took this approach for two reasons. First, we wanted to assess TPs with an open question method (for the reasons given above in the introduction), but it would have been too exhausting and time-consuming for teachers of larger classes to give open-ended descriptions of all their students. Second, we were in any case interested in making inferences about teacher-level effects, so it was more important to get a detailed and valid measure of a teacher's general attitude to his or her class than to gather more superficial information for individual students. There is a similar precedent for this approach in the work of Brophy and Good (e.g., 1970). They also gave feedback to teachers based on observations of only a few students rather than all of them. This approach enabled us to estimate a measure of each teacher's general tendencies, and then use this measure to predict changes in student outcomes for the whole class.

However, within a simple correlational design, it is hard to distinguish effects of TPs on students' outcomes from effects of student outcomes on TPs, because this design does not allow inferences about the direction of causation. If TPs have a causal influence on student outcomes, then they should remain predictive of student outcomes even after controlling for students' actual status at the time the teachers formed their perceptions. In order to assess this, a prospective research design is necessary, in which TPs are measured early in the school year and are then used to predict changes in student outcomes over the course of the school year, having controlled for initial status at the beginning of the year (Jussim 2012).

Therefore, we investigated the relationship of the valence of TPs to student educational attainment, using a prospective research design with two measurement time points and multilevel data analyses. We tested the following hypotheses: The more positive TPs are, (a) the better will be students' future achievement, (b) the more positively students will evaluate their relationship with their teacher, (c) the stronger will be their future learning goal orientation, and (d) the weaker will be their future performance-avoidance goal orientation. Correspondingly, the more negative TPs are, (a) the worse will be students' future achievement, (b) the less positively students will evaluate their relationship with their teacher, (c) the weaker will be their future learning goal orientation, and (d) the stronger will be their future performance-avoidance goal orientation.

3 Method

3.1 Design and samples

At point T1 (2009), at the beginning of Grade 5, we assessed teachers' cognitions and student outcomes. At T2 (2010), on average four months later, we again assessed student outcomes.

In most German federal states, students move after Grade 4 from elementary to secondary school, and this implies a change of teachers. Therefore, teachers and students from Grade 5 were selected to keep the duration of acquaintance between teacher and students constant. We recruited only teachers of mathematics or German (and their students), because these subjects are major subjects in the German school system and therefore contribute decisively to students' educational attainment. We applied a subject-specific approach, to ensure that associations between teachers and students could be uniquely attributed to the respective teacher's perceptions. According to national and state-level statistics (Bildung 2009; Kultusministerium Sachsen-Anhalt 2015), the part of Germany in which we conducted our study is known for a low proportion of male secondary school teachers; the proportion in 2009/2010 was 21.8% (vocational and intermediate tracks). Given this limited availability of male teachers, we decided to study female teachers only, because previous research has already demonstrated that female and male teachers apply different criteria when describing students (e.g., Mullola et al. 2011; Li 1999). For example, female teachers rate girls' persistence and educational competence higher than do male teachers (Mullola et al. 2011). Therefore, we studied female German and mathematics teachers and their students, and assessed all student outcomes specifically for mathematics or German.

We analyzed data from 43 Grade 5 classes in 22 secondary schools (lower tracks: combined vocational and intermediate track) in Eastern Germany. The 43 classes were composed of 23 mathematics and 20 German teachers, females only, with their assigned students, making a total of 635 students (294 female, 338 male, 3 of unknown sex). Students' average age at T1 was $M=10.60$ years ($SD=0.64$, range=10–13 years). In this sample, the number of participants per class ranged from 8 (42.11% of the students in a class) to 25 (92.59% of the students in a class). For some classes teachers reported many students had fallen ill.

At T1, teachers were on average $M=48.05$ years old ($SD=4.97$, range=38–58 years). Teachers had been working at least 10 years in their profession and can be considered experienced teachers (e.g., Rice 2010).

3.2 Measures

3.2.1 Valence of TPs

Valence of TPs was measured by combining a qualitative assessment with a quantitative approach. Each teacher was instructed to think about three randomly-selected

students from their mathematics/German class (we used the last three students listed on the register). Teachers were asked to describe each student in their own words and to mention everything that crossed their minds about these students. Our approach builds on previous methods (e.g., Givvin et al. 2001; Rubie-Davies 2006) but might also have some limitations. For example, it may be that as well as reflecting general teacher attitudes towards students, this method also reflects particular students' attributes to a certain degree (for a detailed discussion, see the limitation Sect. 5.4.)

These descriptions were then segmented into $n=910$ minimum units of information about all aspects of the students. Phrases not related to any aspects of the students ($n=6$) were omitted from further analyses. The remaining information units ($n=904$) were coded according to their valence into either favorable, unfavorable, or neutral. *Favorable* units ($n=495$, 54.8%) were defined as referring to positively valued behaviors, personality features, and/or activities (e.g., "excellent at mathematics", "friendly nature", and "hard working in drama group"), whereas *unfavorable* units ($n=310$, 34.3%) were defined as describing negatively valued behaviors, personality features, and/or activities (e.g., "very bad at writing", "is behaving aggressively", and "according to his parents, he is lazy at home"). *Neutral* units ($n=99$, 10.9%) comprised neutrally valued adverbs and adjectives when describing students' behaviors, personality features and activities (e.g., "displays average performance", "is quiet and reserved").

Two independent raters selected the relevant student descriptions and coded their valence. In doing so, every information unit ($n=904$) was coded by every rater independently. Their independent codings agreed on 97.3% of the entries. Afterwards, the two raters discussed their classifications of the segments that they had not agreed upon until an agreement was reached.

Finally, we computed two valence scores for each teacher across their three student descriptions: (a) the number of their favorable descriptions, and (b) the number of their unfavorable descriptions. Although teachers were not explicitly asked to mention positive and negative aspects (see also Sect. 4.3), every teacher made at least one positive and at least one negative description. Figure 1 displays the associations of positive and negative TPs in the teachers ($r = -0.22$; $p = 0.16$). On average, the teachers mentioned 11 or 12 positive features ($M = 11.51$, $SD = 6.77$, range = 1–37) and 7 negative features ($M = 7.21$, $SD = 4.17$, range = 1–18). Thus teachers described students significantly more frequently in favorable than in unfavorable terms ($t(42) = 11.16$; $p < .001$). Additionally, we tested for differences in the number of favorable ($t(41) = 1.34$; $p = 0.19$) and unfavorable descriptions ($t(41) = -1.47$; $p = .15$) across the subjects mathematics and German. Because we found no such differences, we considered teachers of both subjects together in the following analyses.

3.2.2 Student educational attainment

Student educational attainment was assessed subject-specifically, that is, when reporting school achievement, motivation, and quality of relationship with

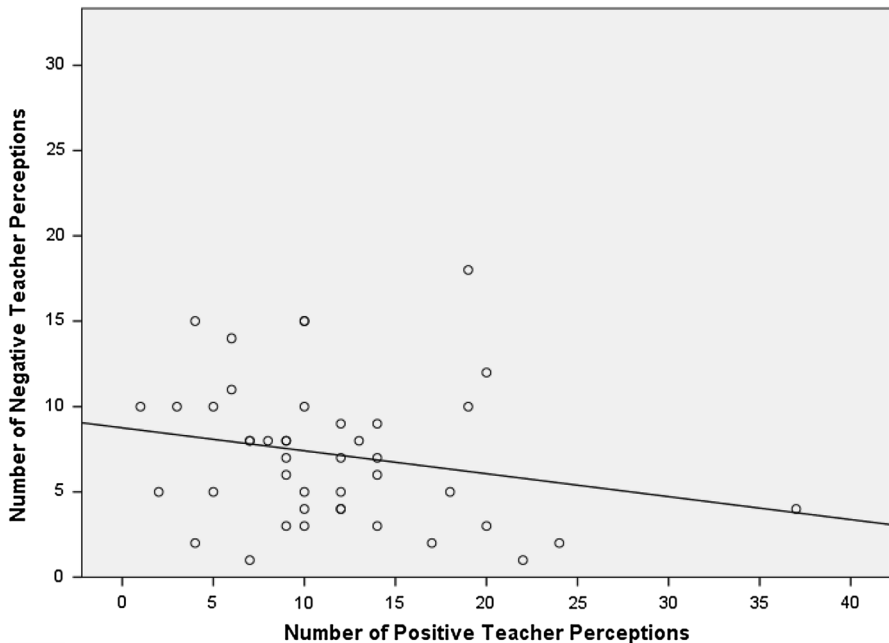


Fig. 1 Distribution of negative and positive valences of teachers perceptions ($N=43$) and their linear association. The valence is based on raters' coding of teachers freely describing three randomly assigned students attending their class

teachers, the instruction and the wording of the items referred to either mathematics or German.

Students' achievement was measured by the grade that the student received in the last major written test of mathematics or German. In the German school system, grades range from 1 ("very good") to 6 ("unsatisfactory"), which is similar to the US with grades ranging from A to F. For ease of interpretation, we reversed these scales so that higher values indicated better achievement.

To measure students' motivation we assessed their performance-avoidance goals and learning goals (see Sect. 1.4). For both goal orientations, we used the Scales for the Assessment of Learning and Performance Motivation (Spinath et al. 2002), a German adaptation of Elliot's Achievement Goal Questionnaire (1999; Elliot and Church 1997). Eight items assessed *performance-avoidance goals*, indicating the motivation not to disclose low competency (e.g., "In mathematics/German class it is important to me not to give wrong answers to questions by the teacher."), eight items measured *learning goals*, indicating the motivation to develop one's competence in terms of task mastery (e.g., "In mathematics/German class it is important for me to learn as much as possible.").

Students' evaluation of the quality of their relationship with their teacher (*student-teacher relationship*) was assessed by the subscale Caring of the Teacher from the Landauer Scales for Social Climate (Saldern and Littig 1987; e.g., "The mathematics/German teacher helps us as a friend."). The scale consisted of eight items.

All responses were given on 6-point rating scales ranging from 1 (“totally disagree”) to 6 (“totally agree”). Scale scores were created by averaging responses across items.

Descriptive statistics, reliabilities and intercorrelations for all T1 and T2 variables are reported in Table 1.

Due to missing data at T2, we used multiple imputation to overcome possible problems concerning biased parameter estimates and reduced statistical power (Peugh and Enders 2004). The average proportion of missing data per variable was 14.80%. We applied multiple imputation by chained equations using the R package mice 3.0.0 (van Buuren 2018). In line with van Buuren and Groothuis-Oudshoorn (2010), we created $m=5$ complete data sets that were then pooled to compute an overall estimate.

4 Results

4.1 Data-analytic approach

We ran separate multilevel models for each indicator of student educational attainment. We applied multilevel modeling because of the nested structure of the educational context (see Sect. 2 for our justification of this approach). At Level 1, we did not enter teacher variables, but at Level 2 valence of TPs was entered to explain differences in student outcomes across classes or schools.

4.1.1 Level-1 analyses

To test the associations between TPs’ valence and student educational attainment, we ran multilevel random-effect models using HLM7 (Raudenbush et al. 2010). We tested the hierarchical structure described above by running null models first. Null models included only random intercepts, which were allowed to vary across the units in the particular model, and no other predictors at any level (Bickel 2007). Three different null models were possible given the structure of our data: (1) a 2-level model with students at Level 1 and classes at Level 2, (2) a 3-level model with students at Level 1, classes at Level 2, and schools at Level 3, and (3) a 2-level model with students at Level 1 and schools at Level 2. Y_{ij} represents the observed outcome for student i in class j and is modeled as the sum of the random intercept for each student in each class, γ_{00} , class-specific variation around this value, u_{0j} , and a level-one residual, ε_{ij} , which refers to the difference between observed and predicted outcomes for a specific student located in a specific class (Model 1). Y_{ijk} refers to student i in class j and school k , whose outcome is modeled as the sum of the random intercept for each student in each class in each school, γ_{000} , school-specific variation around this value, u_{00k} , class-specific variation around this value, u_{0jk} , and the residual ε_{ijk} (Model 2). Y_{ik} represents the outcome of student i in school k modeled as the sum of the random intercept for each student in each school, γ_{00} , school-specific variation around this value, u_{0k} , and the level-one residual, ε_{ik} (Model 3). The models were formulated as follows:

Table 1 Means, standard deviations, reliabilities, stabilities, and intercorrelations among student variables

	<i>M (SD)</i>		α		Measures at T2					
	T1	T2	T1	T2	1	2	3	4	5	6
	1. Sex	0.47				1	−0.01	0.07	0.00	0.05
2. Age	10.60 (0.64)	11.05 (0.65)			−0.07	.54***	−0.19***	−0.04	−0.09*	0.05
3. Achievement	4.08 (1.15)	4.13 (1.16)			0.07	0.08*	.29***	0.11**	0.17***	−0.08*
4. S-T relationship	4.85 (0.89)	4.58 (0.98)	0.71	0.73	0.06	0.00	0.07	.46***	0.36***	−0.12**
5. Learning GO	4.77 (0.81)	4.61 (0.85)	0.77	0.76	0.04	−0.02	0.10*	0.36***	.46***	−0.05
6. Perf.-Avoid. GO	3.31 (1.27)	3.05 (1.23)	0.82	0.83	−0.05	0.03	−0.04	−0.05	0.06	.52***

N = 635 students. Intercorrelations among student variables at T1 are presented below the diagonal, at T2 above the diagonal, and stabilities of student variables are presented along the diagonal. GO: goal orientation, S-T: quality of student-teacher relationship, Perf.-avoid. GO: Performance-avoidance goal orientation; for sex: 0 = male, 1 = female; age is given in years; scale values range from 1 to 6, with higher scores indicating stronger endorsement of the construct

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

$$Y_{ij} = \gamma_{00} + u_{0j} + \varepsilon_{ij} \quad (1)$$

$$Y_{ijk} = \gamma_{000} + u_{00k} + u_{0jk} + \varepsilon_{ijk} \quad (2)$$

$$Y_{ik} = \gamma_{00} + u_{0k} + \varepsilon_{ik} \quad (3)$$

To find the best fitting multilevel model for student educational attainment, we tested these models (1–3) separately for each criterion.

4.1.2 Level-2 analyses

Finally, we entered TP valence as predictors in the best-fitting multilevel-model to explain variance across the units of interest (classes and/or schools). We first centered these valence scores around the grand mean. Effect sizes were determined by calculating local effect sizes, which in our case refers to the proportional reduction in the variance of level-2 intercepts when valence of TPs was added (i.e., the proportion of variance in student outcomes that can be explained by TPs). We then calculated correlation coefficients from this amount of explained variance (Nezlek 2001). In order to test for a directional effect of TPs on educational attainment within the prospective design, we ran a further three models for each criterion. Model 4 tested for the immediate association with educational attainment (at T1), Y_{ij} represented outcomes at T1 of student i nested in unit j predicted by a random intercept for each student in each unit, γ_{00} , unit-specific variation around this value, u_{0j} , the slope for the level-two independent variable teacher perceptions, γ_{01} , and a level-one residual, ε_{ij} . All variables were assessed at T1. Model 5 tested for the long-term association with educational attainment (T2), that is, Y_{ij} referred to outcomes at T2 of student i nested in unit j explained by the same coefficients as in Model 4 with TP assessed at T1. Model 6 built on Model 5 and additionally controlled for autoregression in educational attainment from T1 to T2, in order to test for a link between TP and outcomes independently of any influence of initial student outcomes at the time TPs were formed. Here, the outcome of student i nested in unit j at T2 was additionally predicted on level 1 by student's individual outcome at T1, β_{1j} * student outcome (T1). Our equations were formulated as follows:

$$Y_{ij} (T1) = \gamma_{00} + \gamma_{01} * \text{teacher perceptions} + u_{0j} + \varepsilon_{ij} \quad (4)$$

$$Y_{ij} (T2) = \gamma_{00} + \gamma_{01} * \text{teacher perceptions} + u_{0j} + \varepsilon_{ij} \quad (5)$$

$$Y_{ij} (T2) = \gamma_{00} + \beta_{1j} * \text{student outcome (T1)} + \gamma_{01} \cdot \text{teacher perceptions} + u_{0j} + \varepsilon_{ij}. \quad (6)$$

Table 2 Final random effect estimates for student attainment models at T1 and T2 according to variations across classes and schools

Random effect models	T	<i>SD</i>	χ^2	<i>df</i>	<i>p</i>	ICC
<i>Achievement models</i>						
Model 1: classes at level 2	1	0.52	196.58	42	<.001	0.21
	2	0.38	114.41	42	<.001	0.11
<i>Student–teacher relationship models</i>						
Model 1: classes at level 2	1	0.26	100.13	42	<.001	0.08
	2	0.34	127.84	42	<.001	0.12
<i>Learning goal models</i>						
Model 3: schools at level 2	1	0.15	42.27	20	.003	0.04
	2	0.17	45.85	20	.001	0.04
<i>Performance-avoidance goal models</i>						
Model 3: schools at level 2	1	0.30	53.04	20	<.001	0.05
	2	0.36	69.71	20	<.001	0.08

$N_{\text{Level 1}}=633\text{--}635$ students, $N_{\text{Level 2}}=43$ classes for class level models and 22 schools for school level models. *T* time of measurement, *SD* standard deviation, *ICC* intraclass correlation

Significant coefficients are printed in bold

4.2 Level-1 analyses: means (intercepts) and standard deviations of student educational attainment

4.2.1 Student achievement

The fixed effects for the intercepts were at T1: $\beta_0=4.06$, $SE=0.09$, $p<0.001$ and at T2: $\beta_0=4.13$, $SE=0.07$, $p<0.001$, indicating that average achievement among students was above the midpoint of the scale and stable across a period of four months. This is in line with results of previous studies (Helmke 1998) showing that marks in Grade 5 were favorable on average. Next, we tested for the appropriate multilevel model (1–3) by checking if there were significant random effects, that is, whether the intercepts (β_0) varied significantly across classes or/and across schools (Table 2). Our results supported a 2-level model with students at Level 1 and classes at Level 2 (Model 1), because both measures of students' achievement (at T1 and T2) varied significantly across classes, but not across schools. This significant variation across classes justifies testing for associations with valence of TPs as Level-2 predictors of student achievement.

4.2.2 Student–teacher relationship

The average student–teacher relationship quality was for T1: $\beta_0=4.85$, $SE=0.05$, $p<0.001$ and for T2: $\beta_0=4.55$, $SE=0.06$, $p<0.001$, indicating a significant decrease across the period of our study, $t(634)=6.97$; $p<0.001$. Both measures varied significantly across classes, but not across schools. Again, Model 1, a

2-level model with students at Level 1 and classes at Level 2, appeared to be the best-fitting model.

4.2.3 Learning goals

Students' average learning-goal orientations were $\beta_0=4.76$, $SE=0.05$, $p<0.001$ at T1 and $\beta_0=4.60$, $SE=0.05$, $p<0.001$ at T2, decreasing significantly across the period of our study, $t(634)=4.81$; $p<0.001$. With learning-goal orientations varying significantly across schools, a 2-level model with students at Level 1 and schools at Level 2 was supported (Model 3). Significant differences were also found among classes, but their significance disappeared when entering schools at level 3 (Model 2) and variation across classes seemed to be redundant.

4.2.4 Performance-avoidance goals

The fixed effects for the intercepts were $\beta_0=3.33$, $SE=0.08$, $p<0.001$ at T1 and $\beta_0=3.05$, $SE=0.09$, $p<0.001$, at T2, showing a significant decrease across the period of our study, $t(634)=5.46$; $p<0.001$. Analyses of variances at the different levels supported Model 3, a 2-level model with students at Level 1 and schools at Level 2: Performance-avoidance goals varied significantly across schools. Significant variations across classes were also found (supporting Model 1), but their significance vanished when entering schools at level 3 (Model 2). Thus, variation across classes seemed to be redundant.

In sum, our Level-1 analyses revealed that for each student outcome and at each time of measurement, a 2-level model was the best fitting model. That is, a 2-level model with classes at Level 2 for student achievement and student–teacher relationship, and a 2-level model with schools at Level 2 for students' learning and performance-avoidance goals.

For a variety of student outcomes it has been found that they differ in mathematics and German (e.g., Jurik et al. 2015). Because our classes are nested in subjects, we treated subject as a higher level than class. Thus, we ran 3-level models with students at Level 1, classes at Level 2, and subjects at Level 3 for each criterion and each time of measurement separately. No significant variation across subjects at Level 3 was found for student outcomes at T1 and at T2 (all $ps>0.05$), whereas class level (Level 2) variations remained significant in all models.

4.3 Level-2 analyses: association of TP valence with student educational attainment

Next, we tested whether and which valence of TPs accounted for differences in student educational attainment across classes or schools, that is, we entered TP valences as Level-2 predictors to explain variation in Level-1 intercepts. Separate analyses were run for all T1 and T2 educational attainment models (4–6), as well as for each of the TP valence scores. Results are reported in Table 3.

Table 3 Positive and negative valence of teacher cognitions predicting immediate (T1) and prospective (T2) student attainment using multilevel modeling

Multilevel models	T1 models (Model 4)		T2 models (Model 5)		T2 autoregression models (Model 6)	
	γ (SE)	<i>p</i>	γ (SE)	<i>p</i>	γ (SE)	<i>p</i>
<i>Achievement model</i>						
Positive valence: class level	– .09 (0.07)	.200	.12 (0.04)	.004	.17 (0.07)	.027
Negative valence: class level	– .10 (0.09)	.262	– .04 (0.06)	.503		
Achievement T1: student level					.34 (0.04)	.000
<i>Student–teacher relationship model</i>						
Positive valence: class level	.08 (0.03)	.015	.04 (0.05)	.412		
Negative valence: class level	– .13 (0.04)	.002	– .11 (0.06)	.055		
<i>Learning goal model</i>						
Positive valence: school level	.06 (0.05)	.272	.05 (0.06)	.415		
Negative valence: school level	– .09 (0.04)	.032	– .12 (0.04)	.002	– .08 (0.04)	.045
Learning goal T1: student level					.47 (0.04)	.000
<i>Performance-avoidance goal model</i>						
Positive valence	– .13 (0.07)	.105	– .09 (0.07)	.203		
Negative valence	.09 (0.05)	.091	.18 (0.05)	.002	.14 (0.04)	.002
Perf.-avoid. goal T1: student level					.49 (0.03)	.000

$N_{\text{Level 1}}=633\text{--}635$ students, $N_{\text{Level 2}}=43$ classes for class level models and 22 schools for school level models. Perf.-avoid. goal: Performance-avoidance goal orientation

Significant coefficients are printed in bold

Additionally, we ran analyses including both positive and negative valence scores as Level-2 predictors in the same model.

Because TP valence scores were *z*-standardized, the Level-2 coefficients in Table 3 indicate the change in the Level-1 intercept (i.e., the average student outcome) when the Level-2 predictor (i.e., TP valence score) increases by one standard deviation. For example: The Level-1 intercept for learning goal orientation at T1 was $\beta_0=4.76$. In a 2-level model with schools at Level 2 and negative TP valence as a Level-2 predictor, the Level-2 coefficient for negative TP valence was $\gamma=-0.09$. Therefore, in schools where teachers tend to form 1 standard deviation more negative perceptions than average, the average learning goal orientation of students is 4.67 ($4.76-0.09$). At schools where teachers tend to form 1 standard deviation fewer negative perceptions than average, the average learning goal orientation of students is 4.85 ($4.76+0.09$).

Finally, we tested our Model 6, whether TP valence also contributed to student educational attainment at T2 after controlling for educational attainment at T1 (autoregression; Table 3). To test this model, we entered mean-centered educational attainment at T1 as an additional predictor in all T2 models.

4.3.1 Positive valence of TPs

Positive valence of TPs was unrelated to students' current achievement but explained students' future achievement: More favorable perceptions predicted better student achievement after four months (Model 5, Table 3). This amounted to a variance reduction of 5.82%, corresponding to a correlation coefficient of 0.24. Positive valence was even related to changes in students' performance four months later when controlling for students' achievement at T1 (Model 6, Table 3): The more favorably teachers evaluated their students, the more students' achievement increased after four months. Here, reduced variance was 8.01%, corresponding to a correlation of 0.28. If positive and negative valence scores were simultaneously entered in these analyses, positive valence predicted student performance at T2 significantly ($\gamma=0.12$, $p=0.011$), whereas negative valence did not ($\gamma=-0.01$, $p=0.843$). When both valences were entered simultaneously as predictors of the change in student performance (i.e., in the model controlling for performance at T1), the prediction of positive valence remained significant ($\gamma=0.17$, $p=0.004$), but not that of negative valence.

More positive TPs were also associated with higher *student–teacher relationship* quality at T1 (Model 4, Table 3), but the number of positive TPs did not predict future student–teacher relationship quality. Accounting for positive TPs reduced variation in current student–teacher relationship by 6.21%, equivalent to a correlation coefficient of 0.26. The number of negative TPs was also associated with lower student–teacher relationship at T1 (Model 4, Table 3), accounting for a reduction in variance of 22.61% and an equivalent correlation coefficient of 0.48. The number of negative TPs did not correlate with student–teacher relationship quality at T2. When simultaneously testing the associations of both valences, negative valence predicted the student–teacher relationship quality significantly ($\gamma=-0.12$, $p=0.003$), whereas positive valence did not ($\gamma=0.05$, $p=0.099$).

Overall, positive TPs operated as a class level predictor; explaining only differences across students from different classes. Also, the number of positive TPs was not related to students' motivation.

4.3.2 Negative valence of TPs

In contrast, the number of negative TPs turned out to be a robust predictor of motivational differences between students from different schools. The more negatively teachers from a specific school described their students, the less *learning-goal oriented* were students from this school, currently and prospectively (Model 4 and 5, Table 3). Reductions in variance were 43.41% and 62.65%, respectively, equivalent to correlations of 0.66 and 0.79. Negative valence was even related to changes in students' learning-goal orientation four months later when controlling for students' learning-goal orientation at T1 (Model 6, Table 3): The more unfavorably teachers evaluated their students, the more students' learning-goal orientation increased after four months. This amounted to a variance reduction of 32.29%, corresponding to a correlation coefficient of 0.57.

Similarly robust results were obtained when predicting students' performance-avoidance goal orientation from the number of unfavorable TPs. The more negatively teachers at a school described their students, the more *performance-avoidance goal oriented* students at this school were. This was true for future performance-avoidance goal orientation (Model 5, Table 3). Reduction in variance was 29.55%, equivalent to a correlation of 0.54. The predictive power of negative TPs remained when controlling for students' performance-avoidance goal orientation at T1: Teachers' negative perceptions of students significantly increased students' performance-avoidance goal orientation after four months (Model 6, Table 3). This amounted to a variance reduction of 37.79%, corresponding to a correlation coefficient of 0.62.

We ran additional models to check whether negative TPs were associated uniquely with goal orientations. If the number of positive and negative TPs were included as Level-2 predictors simultaneously, the associations with positive valence (γ s between -0.11 and 0.04) fell below levels of statistical significance in all models (p s between 0.12 and 0.93).¹

5 Discussion

We found that the valence of teachers' perceptions of students in general, an overall measure of their expectations, was associated with various student attributes. Our main findings were that positive TPs predicted students' future achievement and current student–teacher relationship quality. Moreover, negative TPs were associated with lower motivation, both currently and in the future, specifically with lower learning goal orientation and greater performance-avoidance goal orientation. These teacher-related changes in students' achievement and goal orientations occurred over a time period of four months and confirm the results of a small body of research (e.g., over 10 weeks, Trouilloud et al. 2002) showing that the effects of teacher expectations on students can occur over a shorter time period than previously found. (Most studies of naturally occurring teacher expectation effects (e.g., Jussim and Eccles 1992; Rubie-Davies 2006) have been conducted over the time period of a year, measuring teacher expectations at the beginning of the school year and student outcomes at the beginning and end of the school year.) However, most studies of naturally occurring teacher expectation effects (e.g., Jussim and Eccles 1992; Rubie-Davies 2006) have been conducted over the time period of a year, measuring teacher expectations at the beginning of the school year and student outcomes at the beginning and end of the school year.

¹ We additionally performed our multilevel analyses with a combined score, the difference between the number of positive and negative valence statements. When entering the difference as a single predictor in the conditional models (Models 4, 5, and 6), it did significantly explain variance in the outcome (Table ESM A). However, when additionally entering the number of positive or negative valence statements, this significance vanished and the absolute number of entries appears to have more predictive power in all cases (Table ESM B). Results (ESM A and ESM B) can be obtained in the electronic online-supplement (<https://osf.io/dytjq/>).

Positive and negative TPs predicted different educational attainment variables and operated at different levels of the school system: Positive TPs explained differences among students between classes and negative TPs explained differences among students between schools.

5.1 Associations with favorable perceptions by teachers

Positive TPs at the beginning of Grade 5 were unrelated to students' immediate achievement, but significantly associated with students' performance four months later. Because the number of positive TPs predicted change in student achievement, positive TPs can be interpreted as contributing to student outcomes – either as a self-fulfilling prophecy, producing a real change in student performance, or as a positive bias resulting in teachers' awarding better grades to those students that the teachers perceive as better students independently of objective achievement (Jussim 2012). Because positive cognitions were associated only with future student performance and not with current performance, a self-fulfilling prophecy seems to be the more plausible of these two explanations, but this question cannot be answered decisively on the basis of the present study.

This finding agrees with results from class-level studies that found that high expectation teachers' positive expectations of students' future achievement were unrelated to students' current achievement but were related to achievement at the end of the year (Rubie-Davies 2006). The agreement between our results and those from class-level methods suggests that our measure, the number of positive descriptions teachers make of three randomly selected students, is associated with student achievement in the same way as explicit measures of teacher-centered evaluations. Therefore, we will turn to some findings from teacher-centered studies in order to interpret our results.

The positive association between positive TPs and student achievement persisted even after controlling for achievement at the time perceptions were formed, and it was much stronger than the negative association with negative TPs. This is consistent with previous findings from individual-level studies, which found that positive teacher expectations improved student achievement more than negative expectations harmed it (Madon et al. 1997). Our finding is also in agreement with findings on class-level effects, where the positive effects of high expectation teachers on student achievement (d s from 0.50 to 1.44) outweighed the effects of low expectation teachers (d s from -0.03 to 0.20), which were found to be zero or rather small (Rubie-Davies 2006). The explanation given for this pattern is that positive teacher cognitions result in positively encouraging classroom behaviors, whereas negative teacher cognitions do not lead to any particularly consequential classroom behaviors (Rubie-Davies 2006, 2007).

Also, a higher number of positive TPs at the beginning of Grade 5 was associated with a better student–teacher relationship at the time, but not four months later. More specifically, the positive cognitions that teachers reported when describing three randomly selected students reflected how positively students of a given class evaluated teachers' behavior towards them on average. Because this association was

measured at a time when teachers were still getting to know their students, there may be a very general mechanism at work. Teachers might have treated all or most students according to their general level of positive perceptions rather than according to students' individual achievement levels or behaviors.

Our multilevel analyses revealed systematic variation in student achievement and students' evaluation of their relationship to their teacher across classes but not across schools. This appears reasonable. Student achievement was assessed via grades and not standardized tests, thus incorporating teacher perceptions to a certain degree. The student–teacher relationship probably depends on the immediate supportive behavior of the teacher. Because no additional variation was accounted for by schools, interventions to improve these student outcomes should be based on class-level factors (e.g., teacher attributes). Further, when parents want to influence their children's educational success and the socio-emotional climate they experience in classrooms, they might be better advised to try to place their child into a particular class than into a particular school. Nevertheless, when inferring practical consequences, the limitations of our study need to be acknowledged (see Sect. 5.4).

5.2 Associations with unfavorable perceptions by teachers

Negative TPs seemed to be important for students' learning-goal and performance-avoidance goal orientations; negative TPs predicted a significant decrease in students' motivation. Negative TPs explained motivational differences between students attending different schools, and confirmed findings by Wang and Eccles (2013) that students' perceptions of school-environmental factors (e.g., teacher support) prospectively predict student motivation.

These associations might have been mediated by students' sense of competence and of autonomy, because they were related to higher learning (approach) and lower performance-avoidance goal orientations (Cury et al. 2006). Low expectation teachers have been found to provide fewer instructions and explanations about the topic, take less time to relate the topic to students' prior knowledge, and rarely provide regular feedback on students' learning progress (Rubie-Davies 2007); these teacher characteristics might not foster students' sense of competence. Students' learning-approach goal orientation can be understood as part of the intrinsic motivation system (Elliot and Church 1997). Intrinsic motivation can be diminished by asserting external control and undermining the sense of autonomy (Deci and Ryan 2000). Studies of teacher-centered expectations (Rubie-Davies 2007) have suggested that low expectation teachers behave in a manner that does not support autonomy. They group their students in fixed ability-based groups, ask rather closed questions, do not provide students with much scope to interpret teacher questions according to their own understanding, and assert control over students' classroom behavior by procedural comments. However, whether findings from low expectation teachers are fully applicable to our data is a question that needs to be addressed in follow-up studies.

Our finding that students' learning-goal and performance-avoidance goal orientations varied across schools but not classes can also be explained by students' sense of autonomy and competence. Students' sense of autonomy has been shown to be

facilitated by teachers' sense of autonomy (Roth et al. 2007), and teachers' sense of autonomy seems in turn to be related to the school management's level of control (see Sect. 5.3; Eyal and Roth 2011). This makes teachers within a given school more alike than teachers from different schools, which might be reflected in their students' motivation as well.

When the number of positive TPs was additionally entered in the models of motivation, the associations with negative TPs remained unaltered, whereas no associations with positive TPs were found—positive TPs seemed to have no protective effects on students' motivation. Negative TPs reduce students' motivation independently of the number of positive TPs that the teachers also report. One may wonder why positive TPs have no significant association with performance-avoidance and learning-goal orientations. Performance-avoidance goal orientation can be understood as part of the avoidance motive, which is a unipolar dimension selectively susceptible to punishment but not to reward (Fiske et al. 2007), and students may perceive negative and positive TPs as punishments and rewards. If this is the case, then students' performance-avoidance goal orientation should vary with negative TPs but not positive TPs. Alternatively, students' learning-goal orientation can be understood as part of the intrinsic motive system (Elliot and Church 1997), which can be diminished by external punishment, whereas learning-goal orientation is unsusceptible to external rewards (Deci and Ryan 2000). This might explain why students' learning goal orientation is not related to positive TPs.

5.3 Further questions

We did not explicitly consider factors explaining teachers' evaluations. Previous research has already demonstrated negative associations of teachers' evaluations with the school academic composition (an index of students' school difficulties aggregated at school level; Brault et al. 2014). Still, our study implies that school-specific characteristics predict TPs. According to our multilevel analyses, there was significant variance between schools in the number of positive (38.95%) and negative TPs (25.75%). These differences were not due to variations in the type of school (vocational, intermediate, or academic), the age group of the children, or the urban or rural location of the school, since we kept these constant. Brault et al. (2014) reported a positive association between positive school educational climate and teachers' expectations of students. Further, a transactional leadership style of school principals, that is, asserting control by extrinsic rewards and close monitoring of work activities, was found to diminish teachers' sense of autonomy, and increase burnout symptoms (Eyal and Roth 2011; Roth 2014). Conversely, a transformational leadership style, that is, considering the individual teacher, expressing a vision, and fostering intellectual stimulation, correlated positively with teachers' sense of autonomy and negatively with burnout. However, findings on the variation in teacher judgments are correlational and do not allow causal interpretations (Bickel 2007). It is also possible that school-wide student features are associated with teacher cognitions or that third variables influence both student and teacher outcomes. Moreover, most of the variance in the TPs (61.05% for positive TPs and 74.25% for negative

TPs) remained unaccounted for by the different schools, suggesting that other influences might be important. Individual teacher characteristics such as traits might play a role. Research in interpersonal perception (e.g., Funder's Realistic Accuracy Model 2012) demonstrated that people who tend to judge others more positively were described by observers as agreeable, content with life, consistent and not hostile, power-oriented, or anxious (Letzring 2008).

5.4 Limitations

We studied the associations between positive and negative TPs of students in general and various student educational attainment outcomes. We included various sources of information (students, teachers), used multiple methods (quantitative, qualitative), considered the multilevel structure of our data, and investigated our questions in a naturalistic setting using a longitudinal design. However, potential limitations should be considered.

First, as we studied female teachers only, our findings cannot be generalized to male teachers because the genders most likely differ in their TPs (Mullola et al. 2011; Li 1999). The investigated teachers were also experienced teachers and the results should only be transferred with caution to less experienced persons. For example, younger and less experienced teachers seem to apply different criteria when describing students (Bender 1984; Hofer 1986). Further, our participants were enrolled in the German secondary school system and we studied participants from the lower secondary school track (a combination of the vocational and intermediate tracks), in which there are more socioeconomically disadvantaged students (Deutsches PISA-Konsortium 2001, 2010). The school track "could be cautiously interpreted as an indirect indicator for the socioeconomic status of the students" (Peter et al. 2012, p. 62). Further, teachers' impact on students' educational achievement was found to be much stronger in lower secondary school tracks compared to the higher academic track, and the impact of parents much weaker (Deutsches PISA-Konsortium 2001). Similarly, students from lower socioeconomic backgrounds were more sensitive to teachers' expectations than those from higher backgrounds (e.g., Jussim et al. 1996). Thus, our conclusions may be specific to German secondary school students from lower academic tracks. However, this track seems to be practically more relevant for studying associations of teacher perceptions with student outcomes, since students in this track may benefit more from improvements in teacher perceptions.

Second, one may question whether the assessment of TPs via descriptions of just three randomly chosen students actually reflects tendencies of the teachers that are general to their attitude to the whole class and not just to these three students. There are good reasons to think that it does. TPs were assessed when teachers were still getting acquainted with the students, making it less likely that teachers would make highly specific individual assessments. Further, the observed association between TPs of specific students and the average educational outcome of a whole class or school also supports the validity of our approach. Similarly, Rubie-Davies (2006) demonstrated that teacher expectations tended to be teacher-centered rather than depending on attributes of particular students or student groups. However, a higher

number of student descriptions should be included in future studies in order to get more profound information about teachers' cognitions (e.g., six students; Givvin et al. 2001).

Third, student achievement was assessed using grades, which may be more susceptible to teacher bias than standardized tests or end-of-year grades. Thus, the relationships between student achievements and valence of TPs need to be considered cautiously.

Fourth, we used as predictors the absolute numbers of positive and negative perceptions that teachers expressed, and not the proportions of positive and negative statements to the overall number of statements made. Consequently, the numbers of positive and negative statements are both confounded with a teacher's overall tendency to express themselves in detail. However, we think there are good reasons to prefer this measure to a proportion. First of all, the absolute number of teacher perceptions is the "thin slice of behavior" (e.g., Brunswik 1956; Funder 1995) of interest, in the sense that it is the original form in which teachers expressed the valence of their perceptions. Also, research has shown that the valence of teacher expectations is associated with teachers' expressiveness (teachers' reactions to correct answers, teachers' classroom behavior management; Rubie-Davies 2006, 2007), and it may make a big difference for students in the classroom whether a teacher expresses their positive or negative expectations only rarely or extremely frequently. Using absolute numbers of statements allows us to capture this expressiveness, whereas proportional scores would obscure it. For example, if teacher A expresses one positive and one negative perception of a student, and teacher B expresses 10 positive and 10 negative perceptions, this would result in the same proportional scores for both teachers, whereas the absolute score would reflect the fact that teacher A tends to express their positive (and negative) perceptions a lot.

Fifth, our results should be replicated in a larger sample with more Level 2 units. Various Level-2 sample sizes have been recommended, ranging from a minimum number of 20 (Heck and Thomas 2000) or 30 units (Hox 2002), whereas others support the idea of 50 Level-2 units (Bickel 2007) in order to obtain precise estimates.

6 Conclusion

The aim of the present study was to investigate current unknowns in the field of teacher-centered expectations, namely the role of teachers' general way of perceiving and approaching students in daily life. Our study represents an advancement on previous expectancy studies because it demonstrates that teachers' general social cognitions about students also matter for students, that is, when teachers are not explicitly asked to judge particular students' attributes or outcomes.

We looked at the valence of teachers' perceptions of students in general in predicting students' educational attainment. Our results revealed an increase in students' future achievement associated with positive TPs and a decrease in current and future student motivation associated with negative TPs. Assessing the valence of teachers' overall student perceptions offers an alternative approach to understanding the unique role of teacher cognitions in predicting student educational attainment.

Further, positive and negative TPs functioned on different levels: positive TPs accounted for differences among students at the class level and negative TPs at the school level. Given that the valence of TPs appeared to be associated with a variety of student outcomes (achievement, relationship with the teacher, goal orientations), we suggest that the valence of TPs should be addressed in teacher training and considered when discussing school-specific characteristics to improve students' school experience.

Acknowledgements Open Access funding provided by Projekt DEAL.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agirdag, O., Van Avermaet, P., & Van Houtte, M. (2013). School segregation and math achievement: A mixed-method study on the role of self-fulfilling prophecies. *Teachers College Record*, *115*(3), 1–50.
- Agirdag, O., Van Houtte, M., & Van Avermaet, P. (2012). Ethnic school segregation and self-esteem: The role of teacher–pupil relationships. *Urban Education*, *47*, 1135–1159. <https://doi.org/10.1177/0042085912452154>.
- Anderman, L. H., & Anderman, E. M. (1999). Social predictors of changes in students' achievement goal orientations. *Contemporary Educational Psychology*, *24*(1), 21–37. <https://doi.org/10.1006/ceps.1998.0978>.
- Ashton, P. (1985). Motivation and the teachers' sense of efficacy. In C. Ames & R. Ames (Eds.), *Research on motivation in education: The classroom milieu* (pp. 141–174). Orlando, FL: Academic Press.
- Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!*. New York, NY: The Guilford Press.
- Bildung, A. (2009). *Geschlechterdifferenzen im Bildungssystem—Die Bundesländer im Vergleich [Gender differences in the educational system—A comparison on county level]*. München: Druck & Medien Schreiber.
- Brattasani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology*, *76*, 236–247.
- Brault, M. C., Janosz, M., & Archambault, I. (2014). Effects of school composition and school climate on teacher expectations of students: A multilevel analysis. *Teaching and Teacher Education*, *44*, 148–159. <https://doi.org/10.1016/j.tate.2014.08.008>.

- Brophy, J. E. (1985). Teacher–student interaction. In J. B. Dusek (Ed.), *Teacher expectancies* (pp. 303–328). Hillsdale, NJ: Lawrence Erlbaum.
- Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, *61*, 365–374.
- Brunswick, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.
- Butler, R. (2012). Striving to connect: Extending an achievement goal approach to teacher motivation to include relational goals for teaching. *Journal of Educational Psychology*, *104*, 726–742. <https://doi.org/10.1037/a0028613>.
- Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2×2 achievement goal framework. *Journal of Personality and Social Psychology*, *90*, 666–679. <https://doi.org/10.1037/0022-3514.90.4.666>.
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268. https://doi.org/10.1207/S15327965PLI1104_01.
- Demanet, J., & Van Houtte, M. (2012). Teachers' attitudes and students' opposition. School misconduct as a reaction to teachers' diminished effort and affect. *Teaching and Teacher Education*, *28*, 860–869. <https://doi.org/10.1016/j.tate.2012.03.008>.
- Deutsches PISA-Konsortium (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich [PISA 2000: Student literacy in international comparison]*. Opladen, Germany: Leske+Budrich.
- Deutsches PISA-Konsortium. (2010). *PISA 2000. Bilanz nach einem Jahrzehnt [PISA 2009. Résumé after one decade]*. Münster: Waxmann.
- Dusek, J. B., & Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology*, *75*, 327–346. <https://doi.org/10.1037/0022-0663.75.3.327>.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273. <https://doi.org/10.1037/0033-295X.95.2.256>.
- Elliot, A. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*, 169–189. https://doi.org/10.1207/s15326985ep3403_3.
- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*, 218–232. <https://doi.org/10.1037/0022-3514.72.1.218>.
- Elliot, A. J., & McGregor, H. A. (2001). A 2×2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519. <https://doi.org/10.1037/0022-3514.80.3.501>.
- Eyal, O., & Roth, G. (2011). Principals' leadership and teachers' motivation: Self-determination theory analysis. *Journal of Educational Administration*, *49*, 256–275. <https://doi.org/10.1108/0957823111129055>.
- Finney, S. J., Pieper, S. L., & Barron, K. E. (2004). Examining the psychometric properties of the achievement goal questionnaire in a general academic context. *Educational and Psychological Measurement*, *64*, 365–382. <https://doi.org/10.1177/0013164403258465>.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Science*, *11*, 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652–670. <https://doi.org/10.1037/0033-295X.102.4.652>.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*(3), 177–182.
- Givvin, K. B., Stipek, D. J., Salmon, J. M., & MacGyvers, V. L. (2001). In the eyes of the beholder: Students' and teachers' judgments of students' motivation. *Teaching and Teacher Education*, *17*, 321–331. [https://doi.org/10.1016/S0742-051X\(00\)00060-3](https://doi.org/10.1016/S0742-051X(00)00060-3).
- Hamilton, D. L. (2005). *Social cognition: Key readings*. New York, NY: Psychology Press.
- Harris, M. J., & Rosenthal, R. (1985). The mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*, 363–386. <https://doi.org/10.1037/0033-2909.97.3.363>.
- Heck, R. H., & Thomas, S. L. (2000). *An introduction to multilevel modeling techniques*. Mahwah, NJ: Erlbaum.
- Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitsselbstkonzeptes vom Kindergarten bis zur 6. Klassenstufe [From optimists to realists? About the development of the academic self-concept from preschool to 6th grade]. In Weinert, F. E. (Eds.), *Entwicklung im Kindesalter* (pp. 115–132). Weinheim, Germany: Psychologie Verlags Union.

- Hofer, M. (1986). *Sozialpsychologie erzieherischen Handelns. [Social psychology in educational activities]*. Göttingen: Verlag für Psychologie.
- Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.
- Jurik, V., Häusler, J., Stubben, S., & Seidel, T. (2015). Interaction—Erste Ergebnisse einer vergleichenden Videostudie im Deutsch- und Mathematikunterricht [Interaction—First results of a comparative video study in language arts and mathematics instruction]. *Zeitschrift für Pädagogik*, *61*, 692–711.
- Jussim, L. (2012). *Social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy*. New York, NY: Oxford University Press.
- Jussim, L., & Eccles, J. S. (1992). Teacher expectations: II. Construction and reflection of student achievement. *Journal of Personality and Social Psychology*, *63*(6), 947–961. <https://doi.org/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. *Advances in Experimental Social Psychology*, *28*, 281–388. [https://doi.org/10.1016/S0065-2601\(08\)60240-3](https://doi.org/10.1016/S0065-2601(08)60240-3).
- Kuklinski, M. R., & Weinstein R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development*, *72*(5), 1554–1578.
- Kultusministerium des Landes Sachsen-Anhalt. (2015). *Bildungsbericht 2015 Sachsen-Anhalt [Report on education 2015 Saxony-Anhalt]*. Lutherstadt Wittenberg: Elbe Druckerei Wittenberg GmbH.
- Lazarides, R., Buchholz, J., & Rubach, C. (2018). Teacher enthusiasm and self-efficacy, student-perceived mastery goal orientation, and student motivation in mathematics classrooms. *Teaching and Teacher Education*, *69*, 1–10. <https://doi.org/10.1016/j.tate.2017.08.017>.
- Letzring, T. D. (2008). The good judge of personality: Characteristics, behaviors, and observer accuracy. *Journal of Research in Personality*, *42*, 914–932. <https://doi.org/10.1016/j.jrp.2007.12.003>.
- Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*, 63–76. <https://doi.org/10.1080/0013188990410106>.
- Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, *72*, 791–809. <https://doi.org/10.1037/0022-3514.72.4.791>.
- Merton, R. (1948). The self-fulfilling prophecy. *Antioch Review*, *8*, 193–210.
- Mullola, S., et al. (2011). Gender differences in teachers' perceptions of students' temperament, educational competence, and teachability. *British Journal of Educational Psychology*, *82*(185), 206. <https://doi.org/10.1111/j.2044-8279.2010.02017.x>.
- Nezlek, J. B. (2001). Multilevel random coefficient analyses of event-and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, *27*, 771–785. <https://doi.org/10.1177/0146167201277001>.
- Peter, F., Kloeckner, N., Dalbert, C., & Radant, M. (2012). Belief in a just world, teacher justice, and student achievement: A multilevel study. *Learning and Individual Differences*, *22*(1), 55–63. <https://doi.org/10.1016/j.lindif.2011.09.011>.
- Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*, 525–556. <https://doi.org/10.3102/00346543074004525>.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2010). *HLM7: Hierarchical linear and nonlinear modelling [Computer software]*. Lincolnwood, IL: Scientific Software International.
- Rausch, T., Karing, C., Dörfler, T., & Artelt, C. (2016). Personality similarity between teachers and their students influences teacher judgement of student achievement. *Educational Psychology*, *36*, 863–878. <https://doi.org/10.1080/01443410.2014.998629>.
- 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. <https://doi.org/10.3102/0002831210374874>.
- Rice, J. K. (2010). *The impact of teacher experience: Examining the evidence and policy implications. Brief, 11*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research.
- Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, *88*, 408–422. <https://doi.org/10.1037/0022-0663.88.3.408>.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York, NY: Holt, Rinehart & Winston.

- Roth, G. (2014). Antecedents and outcomes of teachers' autonomous motivation. In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Theory and practice* (pp. 36–51). New York, NY: Routledge.
- Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Perceived autonomy in teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology, 99*, 761–774. <https://doi.org/10.1037/0022-0663.99.4.761>.
- Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: Exploring relationships. *Psychology in the Schools, 43*, 537–552. <https://doi.org/10.1002/pits.20169>.
- Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high and low expectation teachers. *British Journal of Educational Psychology, 77*, 289–306. <https://doi.org/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*, 121–135. <https://doi.org/10.1348/000709909X466334>.
- Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2002). *Die Skalen zur Erfassung von Lern- und Leistungsmotivation (SELLMO) [Scales for the assessment of learning and performance motivation]*. Göttingen: Hogrefe.
- Trouilloud, D. O., Sarrazin, P. G., Martinec, T. J., & Guillet, E. (2002). The influence of teacher expectations on student achievement in physical education classes: Pygmalion revisited. *European Journal of Social Psychology, 32*, 591–607. <https://doi.org/10.1002/ejsp.109>.
- van Buuren, S. V. (2018). Package 'mice'. Retrieved from <https://www.stefvanbuuren.nl/mi/>
- van Buuren, S. V., & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, 1–68.
- von Salder, M., & Littig, K. E. (1987). *Landauer Skalen zum Sozialklima [Landauer Scales for the assessment of social climate]*. Weinheim: Beltz.
- Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*, 12–23. <https://doi.org/10.1016/j.learninstruc.2013.04.002>.
- Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding motivation at school. *Journal of Educational Psychology, 91*, 76–97. <https://doi.org/10.1037/0022-0663.91.1.76>.
- Wirtz, M. A., & Strohmmer, J. (2014). *Dorsch-Lexikon der Psychologie [Dorsch encyclopedia of psychology]*. Bern, Switzerland: Hans Huber Verlag.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dr. Nancy Tandler is a postdoctoral researcher at the Department of Assessment and Differential Psychology at the Martin-Luther-University (MLU) Halle-Wittenberg. Her main interest is the role of individual differences in explaining educational outcomes in students and teachers.

Prof. Dr. Claudia Dalbert was a full professor of Educational Psychological at the Martin-Luther-University (MLU) Halle-Wittenberg. Her main interest is research the justice motive as a personal resource. Currently, she is Minister of Environment, Agriculture and Energy of the federal state of Saxony-Anhalt, Germany.