University of Groningen

The Interpersonal Character of Teacher Expectations

Timmermans, A. C.; van der Werf, M. P. C.; Rubie-Davies, Christine M.

Published in:
Journal of School Psychology

DOI:
[10.1016/j.jsp.2019.02.004](https://doi.org/10.1016/j.jsp.2019.02.004)

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
Final author's version (accepted by publisher, after peer review)

Publication date:
2019

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

Citation for published version (APA):

Timmermans, A. C., van der Werf, M. P. C., & Rubie-Davies, C. M. (2019). The Interpersonal Character of Teacher Expectations: The Perceived Teacher-Student Relationship as Antecedent of Teacher Expectations. *Journal of School Psychology, 73*, 114-130. <https://doi.org/10.1016/j.jsp.2019.02.004>

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.

The Interpersonal Character of Teacher Expectations: The Perceived Teacher-Student
Relationship as Antecedent of Teacher Expectations

A. C. (Anneke) Timmermans¹, M. P. C. (Greetje) van der Werf¹, and Christine M. Rubie-
Davies²

¹University of Groningen, GION Education/Research

²The University of Auckland, Faculty of Education and Social Work

Corresponding author: A. C. Timmermans, University of Groningen, GION
Education/Research, Grote Rozenstraat 3, 9712 TG, Groningen, the Netherlands,
A.C.Timmermans@rug.nl, +31(0)50-3635740

Prof. dr. Greetje van der Werf, University of Groningen/ Groningen Institute for
Educational Research, Grote Rozenstraat 3, 9712 TG Groningen, The Netherlands,
M.P.C.van.der.Werf@rug.nl

Prof. dr. Christine Rubie-Davies, The University of Auckland, Faculty of Education
and Social Work, c.rubie@auckland.ac.nz

This research was approved by the ethical committee of the Pedagogical and
Educational Sciences department of the University of Groningen.

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

The Interpersonal Character of Teacher Expectations: The Perceived Teacher-Student Relationship as Antecedent of Teachers' Track Recommendations

Abstract

Teacher expectations of students have long been recognized as a form of interpersonal expectations. In this study, we aimed to investigate the interpersonal character of teacher expectations by assessing 1) whether teacher expectations and the teacher-student relationship shared similar antecedents in terms of demographic characteristics of students, and 2) whether the dimensions closeness, conflict, and dependency of the teacher-student relationship were predictive of teacher expectations. Analyses were based on a large sample of 9,881 students in 614 classes in the final grade of primary education. The results indicated that teacher expectations – as measured by track recommendations - and the teacher-student relationship were not consistent over antecedents. Student performance and parental education were positive predictors of track recommendations and closeness, whereas they were negatively associated to conflict and dependency. Ethnicity was positively associated to track recommendations, but negatively to closeness. Furthermore, perceived closeness and conflict were not statistically significantly associated with track recommendations. A negative association was found for perceived dependency with teachers' track recommendations, although the latter association appeared stronger for high performing students. Finally, a different weighting of the conflict and dependency applied between teachers and classes when formulating track recommendation, indicating that it did play a (stronger) role in some of the classes.

Keywords: Teacher expectations; teacher-student relationship; elementary schools; multilevel analysis

Introduction

Teacher expectations of students have long been recognized as a form of interpersonal expectations, that is, the beliefs that one person, for example a teacher, has about the future performance or behaviors of another person, for example a student. We may therefore assume that teacher expectations are dependent on multiple sources, including characteristics of the teacher holding the expectation (e.g., Rubie-Davies, 2010; Babad, 2009; Weinstein, 2002), characteristics of individual students who are the subject of the expectations (e.g., Ready & Wright, 2011; Tenenbaum & Ruck, 2007), but also on the quality of the interpersonal dyadic relationship between a teacher and a student as perceived by the teacher holding the expectation (Hughes, Gleason, & Zhang, 2005). Similarly, how the interpersonal dyadic relationship, that is, the unique enduring connection between two individuals (a teacher and a student; e.g., Collins & Repinski, 1994; Hinde, 1997), is perceived might also be affected by the teacher's and student's individual characteristics, as well as how the student is perceived by the teacher. In this study, we aimed to investigate the association between these two interpersonal constructs, by testing (1) whether teacher expectations and teachers' perceived relationships with their students shared similar antecedents in terms of demographic characteristics of individual students, and (2) whether the teacher-student relationships as perceived by the teacher predicted teacher expectations after other characteristics of students, including performance, were taken into account.

The current study was focused on the association between teacher expectations and teacher-student relationships as perceived by teachers in the Dutch context; a context in which teacher expectations play a major role in the transition from primary to secondary education in the form of track recommendations. In the following literature overview we first describe the Dutch context, track recommendations as a special case of expectations, and the importance of track recommendations for the students' educational careers. Thereafter, we

focus on teacher expectations in general, interpersonal relations between teachers and students, and the potential of the teacher-student relationship to be associated with teacher expectations.

The Dutch Educational System and the Importance of Track Recommendations

Dutch primary education is intended for all children from age 4 (prekindergarten) up to, and including, age 12 (Grade 6). Besides regular primary schools, special schools for primary education and schools for special education provide education in smaller classes and have specific expertise for the students who need additional support that cannot be provided by regular primary schools. In 2013, approximately 1,500,000 pupils were enrolled in 6,500 primary schools (Ministry of Education, 2014).

Primary schools are obliged to monitor the progress of their students during primary education by means of monitoring and evaluation systems. These monitoring systems offer schools and teachers the possibility to monitor the progress of their students during primary education via several instruments, such as a set of tests, a registration system, remediation guidance methods, and tools for identifying specific learning problems. In the final grade, about 85% of the schools administer the “Cito School Leavers’ Test”, a highly reliable standardized test which includes the basic subjects: Dutch language, mathematics, and information processing. This test was explicitly designed to assist teachers in formulating a track recommendation for secondary education (van Boxtel, Engelen, & de Wijs, 2011). The value of the test has been debated frequently, most of all because in Dutch media the test outcome is often described as a snap-shot. Nevertheless, the predictive validity of the tests from the monitoring system and the school leavers test for students’ school success – as measured by track level in fourth year in secondary education – are both high (van Aarsen, Roeleveld, & Luyten, 2013).

Before the transition from primary to secondary education, each student receives a track recommendation, which is considered to be an expression of the teacher's expectations for the student's future performance during secondary education (e.g., De Boer, Bosker, & Van der Werf, 2010; Inspectorate of Education, 2007). Teachers need to indicate for each student which track they consider optimal given the students' potential to develop. With this expectation, teachers are encouraged to take into account the students' performance (scores on the school leavers' test and monitoring system tests), the teachers' professional opinion of student behavior, and other factors they deem relevant for success in secondary education (Inspectorate of Education, 2007).

Dutch secondary education is, with six different tracks, a highly tracked system and the tracking takes place at a relatively young age (approximately 12 years). On enrollment, students in Dutch secondary education are placed in a specific track (or sometimes in a class combining two adjacent tracks) based on their scholastic aptitude. Until 2014, placement decisions were based on the combination of track recommendation and score on the school leavers' test, whereas after 2014 the track recommendation became the sole determinant of track placement.

The duration of the tracks varies between four (the lower tracks) and six years (the highest track), and each track offers different access to further education. Besides the differences in level and length of the tracks, they also vary in the content provided to the students. The highest tracks (pre-university education, higher general secondary education, and the theoretical track of pre-vocational education) offer students a general academic preparation for future education in universities, higher professional education or senior secondary professional education. The lowest tracks prepare students for senior secondary vocational education by providing a vocational program in the third and fourth year. Tracking in Dutch secondary education takes place partly within schools and partly between schools.

Some schools offer the full range of tracks while others are single or double track schools. Grade repetition within tracks and intermediate upward or downward mobility between the tracks is possible, as students can change tracks depending on their grades. Relative to their track recommendation, 35% of the students change grades in the first three years of secondary education (Timmermans, Kuyper, & Van der Werf, 2013). Several studies have shown that the teachers' track recommendations have long-term effects during secondary education and beyond (e.g., De Boer et al., 2010, Feron, Schils, & ter Weel, 2016). For example, De Boer et al. (2010) showed that track recommendation bias was highly predictive of initial placement of students in secondary school classes, and continued to be predictive until the fourth grade in secondary education when the students were approximately 16-17 years old. As the Dutch system shows a growing inequity and increasing rigidity in placement and mobility, one of the main challenges for future educational policies is to consider options for reducing tracking and to increase the quality of the track recommendations (OECD, 2016).

Teacher Expectations

Teacher expectations are “primarily cognitive phenomena, inferential judgments that teachers make about probable future achievement and behavior based upon the student's past record and his present achievement and behavior” (Brophy & Good, 1974; p. 129). The expectations teachers have for their students have also been described as inferences made by teachers with respect to students' potential to achieve (Riley & Ungerleider, 2012). Given these definitions, track recommendations, as explained above, may be seen as a special case of teacher expectations and although they are not the most common indicator of teacher expectations, track recommendations have been labeled as teacher expectations or special cases of expectations in several earlier studies (e.g., de Boer, Bosker, & van der Werf, 2010; Glock & Krolack-Schwerdt, 2013; Timmermans, Kuyper, & Van der Werf, 2015).

The importance of teacher expectations in facilitating students' learning has long been recognized as they may create a self-fulfilling prophecy (Rubie-Davies, 2008). A self-fulfilling prophecy can only arise when the initial teacher expectations were inaccurate, but the expectation later becomes confirmed as people tend to interact according to the expectation (Madon, Willard, Gyll, Kyle, & Scherr, 2011; Merton, 1948). In general, teacher expectations influence teacher behavior and the subsequent performance of students, although this seems to depend on a range of teacher behaviors, teacher characteristics, and student characteristics (e.g., Brophy & Good, 1970; Ready & Wright, 2011; Rubie-Davies, 2008, 2010). Relative to low expectation students, teachers demonstrate a positive bias in evaluating the work of high expectation students, provide them with more response opportunities, more challenging instruction, more praise, and interact with them in ways that are more supportive and caring (e.g., Babad, 1992; Jussim & Eccles, 1992; Jussim, Eccles, & Madon, 1996). This differential treatment of high and low expectation students may account, at least partially, for the expectancy-confirming impact of teacher expectations (Hughes, Gleason, & Zhang, 2005). It is very likely that in the short term track recommendations affect teacher behavior (e.g., asking questions, feedback provision, and opportunity to learn) towards the class in general as well as to specific high or low expectation students in a similar way as regular teacher expectations are portrayed. Unfortunately, to our knowledge there is no empirical evidence yet to support this assertion. Whereas regular teacher expectation operationalizations are associated with ability grouping within or across classes (Brophy, 1983), track recommendations have a less subtle effect on subsequent performance of students as they directly influence the track placement decisions. The consequences of track placement or streaming are far-reaching as the level, content, amount of learning materials, and instruction and support differ by track/stream and the track/stream in which students are placed is partially dependent on teacher expectations (e.g., Brophy, 1983; Kelly & Carbonaro,

2012; Metz, 1978). Furthermore, many tracked systems are limited in the possibilities for students to change tracks (Jacob & Thieben, 2007; Lucas, 2001) and hence the original expectation becomes fulfilled via the opportunity to learn provided in the different tracks. This is supported by several studies which have shown that the effects of track recommendations remain visible throughout the educational careers of students (e.g., Caro, Lenkeit, Lehman, & Schwippert, 2009; De Boer et al., 2010).

Another body of studies relates to antecedents of teacher expectations (i.e., the question of whether there are particular subgroups of students for whom teachers are less accurate in their expectations). In these studies, the focus is almost exclusively on non-malleable characteristics of students, and in particular on students' demographic background (gender, socioeconomic status, minority status), assuming that teacher expectations are partly based on stereotypes (Jussim, Eccles, & Madon, 1996). Results from some recent studies have indicated that teachers tend to have lower expectations for the future academic performance of boys, minority group students, and students from less affluent families, after taking the students' performance into account (e.g., Glock & Krolak-Schwerdt, 2013; Timmermans et al., 2015). There also appears to be a slowly growing body of literature examining other characteristics of students, mainly related to student behavior as perceived by the teachers. These studies have indicated that, after controlling for performance levels of students, teacher expectations seem to be positively associated to students' perceived assertiveness, independence, self-confidence, engagement, and the students' social behavior (e.g., Alvrdez & Weinstein, 1999; Bonvin & Genoud, 2006). These studies have indicated that teacher expectations may be derived from multiple sources, such as student performance, stereotypes, and teacher perceptions of student behaviors. Similar antecedents are found when it comes to track recommendations. After taking the students' current performance into account, track recommendations are generally lower for boys, majority students, and students

from less affluent families, and positively relations have been found with students' perceived self-confidence, engagement, and social behavior (Rubie-Davies, 2010; Timmermans et al., 2015; Timmermans, de Boer, & Van der Werf, 2016). Until now, however, the role of interpersonal relations has been largely overlooked in teacher expectation research, in general, as well as for track recommendations specifically.

The Interpersonal Teacher-Student Relationship as Perceived by the Teachers

Relationships are frequently defined as enduring connections between two individuals, uniquely characterized by degrees of continuity, shared history, and interdependent interactions across settings and activities (Collins & Repinski, 1994; Hinde, 1997).

Definitions of interpersonal relationships are frequently extended to include the qualities of a relationship, as indicated by levels of trust, intimacy, and sharing; the presence of positive affect, closeness, and affective tone; and the content and quality of communication (Collins & Repinski, 1994; Laible & Thompson, 2007). From the early 1990s onward, scholars began investigating interpersonal dyadic teacher–student relationships (e.g., Koomen, Verschueren, van Schooten, Jak, & Pianta, 2012; Mantzicopoulos & Neuharth Pritchett, 2003; Pianta & Steinberg, 1992), mostly by measuring how these relationships were perceived by teachers.

Since the study of Pianta, Steinberg, and Rollins (1995), the three-dimensional model of closeness, conflict, and dependency has dominated the teacher–student relationship literature. Within this framework, closeness reflects the degree of openness, warmth, and security in the relationship according to the teacher; conflict refers to the degree to which a teacher perceives teacher–student interactions as negative, discordant, unpredictable, and unpleasant; and dependency denotes a developmentally inappropriate degree of overreliance and possessiveness of the child in the relationship (Pianta, 2001).

Closeness is generally viewed as a positive relational factor, providing children with emotional security and support to deal with the socioemotional and academic demands they

face in school. Conflict and dependency, on the other hand, are perceived as negative relational factors, both reflecting a lack of security, and consequently hampering and interfering with children's coping with demands in the school context (Koomen et al., 2012). Higher levels of closeness as perceived by teachers have, for example, been found to be associated with better cognitive functioning among students and improved task behavior, language and mathematics skills, prosocial behavior, popularity with peers, and with less social withdrawal and aggressive, antisocial, and hyperactive behavior (e.g., Birch & Ladd, 1997; Palermo, Hanish, Martin, Fabes, & Reiser, 2007; Thijs, Koomen, & van der Leij, 2008). Higher levels of conflict, on the other hand, have been shown to be associated to, for example, less classroom participation, more negative school attitudes, less productive future work habits, lower language and mathematics grades, less prosocial behavior and more aggression, disruption, hyperactivity, and externalizing problem behavior, and future disciplinary problems (e.g., Birch & Ladd, 1997, 1998; Doumen, Verschueren, Buyse, Germeijs, Luyckx, & Soenens, 2008; Hamre & Pianta, 2001; Palermo et al., 2007; Pianta et al., 1995; Thijs et al., 2008). Similarly, high levels of dependency, as perceived by teachers, have been found to associate to more negative school attitudes, less prosocial behavior, more social withdrawal and loneliness, more aggression, hyperactivity, and antisocial behavior with peers (e.g., Birch & Ladd, 1997, 1998; Hamre & Pianta, 2001; Palermo et al., 2007; Thijs et al., 2008). The aforementioned associations between teacher perceptions of the student–teacher relationship quality and subsequent adjustment or development usually hold when previous levels of adjustment have been statistically controlled. For example, in the study of Hughes, Cavell, and Jackson (1999), teacher ratings of the student-teacher relationship in Year 1 were positively related to students' aggression, as rated by the teacher in Year 2, after controlling for aggression ratings in Year 1.

Current research has mostly focused on the effects of student–teacher relationships on students’ well-being and performance in various settings. Research on antecedents of the student–teacher relationship quality as perceived by the teacher (i.e., child and teacher characteristics that predict a supportive student–teacher relationship) has, in contrast to research on antecedents of teacher expectations, been primarily focused on students’ perceived behaviour, such as internal or external problem behaviour, temperament, attachment, engagement and motivation (e.g., McGrath & Van Bergen, 2015; Nurmi, 2012). However, differences in student–teacher relationship quality have been reported in association to several non-malleable characteristics of students, including gender, socioeconomic status, and ethnic background (e.g., Birch & Ladd, 1997; Entwisle & Alexander, 1988; Hughes et al., 2005; Koepke & Harkins, 2008; McGrath & Van Bergen, 2015; Murray & Murray, 2004; Saft & Pianta, 2001; Thijs, Westerhof, & Koomen, 2012; Zee & Koomen, 2017). For example, African-American children are less likely to experience supportive relationships with teachers, especially when their teacher is non-African-American (Saft & Pianta, 2001). Additionally, for young children, teachers have reported higher levels of closeness in their relationships with girls compared to boys, with whom relationships are characterized by higher levels of conflict (Birch & Ladd, 1998, Koepke & Harkins, 2008; Silver et al., 2005). However, results are far from consistent, as Murray and Murray (2004), for example, found no significant associations between gender and the dimensions of teacher-student relationships, and for ethnicity only the dependency dimension was found to be significant. As a construct that is explicitly interpreted as interpersonal, and thus presumably dependent on the characteristics of both teachers and students involved and perceptions of one another, research into student-teacher relationships may benefit from a more systematic study of antecedents.

The Potential Association between Teacher Expectations and Teacher-Student Relationships

Although both constructs, teacher expectations and teacher-student relationships, deal with the interpersonal dyadic relations between a teacher and a student, and with the teacher's perception of the student, the concepts have rarely been investigated simultaneously. However, there seems some potential synergies in the association between the teacher-student relationship and subsequent teacher expectations. Given the assumed interpersonal character, we may expect both the teachers' expectations as well as the teachers' perception of the student-teacher relationship to be related to various sources, including teacher and student characteristics (as perceived by the teacher). As the same actors are involved, and both constructs develop within the same educational context, we may expect some similarities in the antecedents of both constructs. More research is however necessary to investigate whether this is actually the case.

Additionally, some scholars have argued that it may be the teacher's perceived teacher-student relationship that is principally influential in educational decisions, the marks given, and actions toward students in instructional and interpersonal interactions (Ang, 2005; Hamre & Pianta, 2001; Thijs et al., 2008; Wentzel, 2012). These studies assume that the perceived teacher-student relationship is a precursor of teacher behavior. Following this line of reasoning, it may be the teachers' perceived relationship with students that serves as a source of information on which teachers build their expectations. In other words, the expectations that a teacher has for particular students may be influenced by how much students are appreciated by their teacher. In particular, teachers may have high expectations for students with whom they have positive relationships. The strength of these relationships then bias teachers' expectations such that they have lower expectations than warranted for students with whom they have difficult and more fractious relations.

Instead of this unidirectional association, some scholars have argued for a reciprocal development of expectations and relationships throughout the school year (e.g., Hughes et al., 2005; Timmermans et al., 2016), because the relationship between teachers and students is considered one of the most important mediators through which expectations affect student outcomes (e.g., Brophy & Good, 1970, Harris & Rosenthal, 1985). However, no study has investigated whether the relationship between these two interpersonal constructs is a reciprocal one. It is important to note that it is the perception of the teacher that is crucial here (Thijs et al., 2008), not the relationship itself, or the student-reported relationship. Therefore, we will refer to the perceived teacher-student relationship to stress the importance of the teachers' perceptions.

In three empirical studies, there have been indications of associations between teacher expectations and the general teacher-student relationship whereby, in all three studies, the interpersonal relationship as perceived by a teacher has been considered as a precursor to teacher expectations. Hughes et al. (2005) found that the general quality of the teacher-student relationship, as perceived by the teacher, was more closely associated to teachers' expectations than to the children's measured performance and background. Controlling for children's actual abilities, parent educational level, child ethnicity, and gender, when teachers viewed their relationships with children as less affectively positive, they rated children as less academically competent. Similarly, Rubie-Davies (2010) found that teachers who had high expectations for all their students also perceived the teacher-student relationship (one-dimensional indicator) more positively than low expectation teachers. Moreover, for high expectation teachers there was a significant positive association between expectations of academic performance and the teacher-student relationship. In the Dutch context, the association between track recommendations and the general teacher-student relationship appeared positive and significant, after taking the performance of students into account

(Timmermans et al., 2016). However, this association disappeared after the students' work habits, as perceived by the teacher, were taken into account. Further, the study showed that the association between expectations and perceived teacher-student relationships varied to a large extent between classes, indicating positive relations in some classes and negative relations in other classes. Common to the three studies described above is that teacher reports of overall measures of the teacher-student relationship were used instead of the dimensions of closeness, conflict, and dependency. More research is necessary to assess whether there seem to be particular dimensions of the teacher-student relationship that are predictive of the teachers' expectations.

The Current Study

The interpersonal nature of teacher expectations is the central focus of the current research. In this study, we aimed to investigate (1) whether teacher expectations and the dimensions of closeness, conflict, and dependency of the teacher-student relationships as reported by teachers shared similar antecedents for children aged 11-12 (gender, ethnic background and parental education), and (2) whether the teachers' expectations were associated with the dimensions of closeness, conflict, and dependency of the teacher-student relationship, after the performance of students was taken into account. We tested the following hypotheses in a large sample of 9,881 students in the final grade of Dutch primary education.

First, we expected gender, parental education, and ethnic background to be antecedents of both teacher expectations and the perceived teacher-student relationship (H1). We chose particularly these antecedents to stay in line with previous teacher expectation research. In addition to the existing body of evidence, we also investigated the interactions between students' background and performance to examine whether there appeared to be particular combinations of student characteristics that were associated with higher or lower

levels of expectations and teacher-student relationship perceptions. Second, we expected the perceived teacher-student relationship to be associated with teacher expectations (H2), because there may be a reciprocal association between these two interpersonal constructs (Hughes et al., 2005). In the current study, due to the temporal design of the data collection, we could only investigate teacher-student relationships as precursor of teacher expectations. With respect to these relationships, we expected a positive association between the dimension of closeness and teacher expectations and a negative association between the dimensions of conflict and dependency with teacher expectations. Similarly to the first hypothesis, we also investigated whether interactions between perceptions of the teacher-student relationship with other student characteristics added to the prediction of the teachers' expectations.

Method

Sample and Procedure

The empirical analyses were based on measurements of the COOL⁵⁻¹⁸ cohort studies conducted in 2007/2008 (Driessen, Mulder, Ledoux, Roeleveld, & van der Veen, 2009). The COOL⁵⁻¹⁸ cohort study is a large scale assessment in the Netherlands with measurements every three years. Students were followed from the age of 5 to approximately the age of 18. The 2007/2008 cohort consisted of students in 400 elementary schools who were representative of Dutch elementary schools, supplemented with a sample of students in 150 schools with relatively high proportions of minority students and schools with alternative educational concepts, such as Waldorf and Montessori schools. In the school year 2007/2008, data were collected in these schools in Kindergarten, Grade 3, and Grade 6, but, for the purpose of this study, we used only the Grade 6 data.

Data collection took place in several steps. First, schools were asked to participate in the Fall of 2007. If schools decided to participate, lists of students and student background information were collected from the schools' administration (Fall 2007 – Winter 2008).

Second, parents gave informed consent. Third (Spring 2008), schools administered the cognitive achievement tests, which was a regular activity for many schools who administer these tests for their own monitoring systems. Concurrently, teachers completed questionnaires concerning the teacher-student relationship for each of their Grade 6 students. Finally, (late Spring 2008) teachers were asked to formulate their recommendations. These steps in data collection ensured that student background, student performance, and perceived teacher-student relationship were all measured before the formulation of the teachers' track recommendations.

The Grade 6 sample consisted of 9881 students in 614 classes of 485 primary schools. Of the participating students, 50.3% were boys and 49.7% were girls, 76% of the students had a Dutch background, whereas 24% had a different background. The largest minority groups in the Netherlands represent Turkish and Moroccan students; they comprised, respectively, 30.9% and 25.5% of the minority students in the current sample. Other, much smaller groups of minority students stem from a variety of backgrounds related to Dutch colonial history (Suriname, the Dutch Antilles), the civil war in Yugoslavia, and instability in Afghanistan, Iran, Iraq, and Somalia (Stevens, Clycq, Timmerman, & van Houtte, 2011). With respect to the level of parental education, the groups of students whose parents had low (primary education and lower secondary education), middle (upper secondary education and senior secondary vocational education), and high (higher professional education or university) parental education consisted of 31.8%, 41.2% and 27.0%, respectively.

Instruments and Variables

Teacher expectations. Teacher expectations were operationalized using the track recommendations given by teachers for their students at the end of elementary education (late Spring or approximately eight months into the school year). This track recommendation is considered an informed expectation of the teacher indicating which track is optimal for a

student given the student's potential according to the teacher. Considering the typology of measures of teachers' expectations of Hoge (1984), this operationalization can be seen as a global estimate of the student's academic potential, a one-dimensional index of teacher expectations. This variable has been used in research before as an operationalization of teacher expectations (e.g., De Boer et al., 2010; Timmermans et al., 2015).

The track recommendations for the students were requested by means of a teacher questionnaire. Teachers were allowed to indicate a single or two adjacent tracks. The recommendations were coded on a scale called the "educational ladder" (Bosker & Van der Velden, 1985, 1989) which was developed to map the positions of students in Dutch secondary education. The original educational ladder has scores ranging from 1 – 12, with the score of twelve representing a student who has graduated from the pre-university track. For every year or track away from graduation in the pre-university track, a point is subtracted. Possible values for track recommendation or entry to secondary education are 1 to 6. The educational ladder was developed and treated as an interval scale because it referred to the number of years to the top of Dutch secondary education. The educational ladder used for this study differed slightly from the original, due to changes in the educational system between the 1980's and 2007 (e.g., De Boer et al., 2010; Roeleveld, Driessen, Ledoux, & Meijer, 2011). In the current study, values between 1.5 and 6 were present, as special needs education was not considered in the sampling. Each step of 1 point corresponds to a school track, for example, 6 refers to the single pre-university track recommendation. When teachers indicated two adjacent tracks, for example higher general secondary education and pre-university track recommendation, the average score on the educational ladder was used. In this example, this equalled 5.5 as the average was of pre-university education (6) and higher general secondary education (5).

Student performance. The students' scores on the school leavers' test, which was developed by Cito, The Netherlands Institute for Educational Measurement, were available in the dataset. The test was taken in early springtime. The test contains three parts, including Dutch language (100 items), mathematics (60 items), and information processing (40 items). Students' scores were constructed by means of the One-Parameter-Logistic Model (OPLM, Verhelst & Glas, 1995) and converted by Cito to a scale ranging from 501 to 550 (van Boxtel et al., 2011). The school leavers' test is a highly reliable ($MAcc = .96$), high-stakes test. The $MAcc$ coefficient (Accuracy of Measurement) is a reliability coefficient with a similar interpretation as classical reliability coefficients, and was retrieved from the IRT-model used to scale the items (Verhelst, Glas, & Verstralen, 1995). In addition, the test is highly predictive of student school success in secondary education, as the explained variance for the position in the fourth year of secondary education is .497 (Van Aarsen et al., 2013). Every year a new school leavers test is developed, which is similar in structure but has different questions to previous tests. An extensive pilot test is conducted every year to ensure that similar performance is awarded similar test scores. The school leavers tests of 1997 and 2010 have been investigated in terms of Differential Item Functioning (DIF) for students with different backgrounds (Van Boxtel et al., 2011; Van Schilt-Mol, 2007), indicating that although the tests contained some items which revealed DIF, overall, the test did not advantage or disadvantage particular student subgroups and was therefore considered culturally unbiased.

In addition, the COOL⁵⁻¹⁸ cohort contains three tests that provided information on the achievement of students during the final year of elementary education, which stem from the compulsory monitoring and evaluation systems. These tests were administered during the winter or early spring. All three tests have been developed and assessed to sufficiently cover the content and goals of the curriculum of Dutch primary education. The language test

consisted of 64 items (MAcc .90; range 9 - 177), the mathematics test of 120 items (MAcc .96; range 54 - 160), and the reading comprehension test of 50 items (MAcc .87; range 1 - 100). All three tests can therefore be considered highly reliable. In addition the tests are highly predictive of students' success in Dutch secondary education, as the explained variance in positions in the fourth grade of secondary education of the combined three tests is .537 (Van Aarsen et al., 2013). The scores used in the analyses were the scores on the underlying latent scales of language, mathematics, and reading comprehension (Janssen, Verhelst, Engelen, & Scheltens, 2010; van Berkel, Hilte, Groenen, & Engelen, 2014; Weekers, Groenen, Kleintjes, & Feenstra, 2011). These scores were derived from the One-Parameter-Logistic Model (Verhelst & Glas, 1995).

Student background. Student background information was retrieved from the schools' administration. Student background variables were included in the analyses because previous research in the Netherlands has shown differences between these groups in track recommendations (De Boer et al., 2010; Timmermans et al., 2015).

A dummy variable was created for gender, where boys formed the reference group. Parental education was an ordinal variable with three categories, including the levels low (primary education and lower secondary education, ISCED level 1 - 2), middle (upper secondary education and senior secondary vocational education, ISCED level 3 - 5), and high (higher professional education or university, ISCED level 6 - 8). Two categories represented the ethnic background of the students (Dutch and students with a minority background), which was based on information concerning the birth countries of the students' parents.

Teacher-student relationship: closeness, conflict, and dependency. For every student in their class, teachers answered 15 questions concerning the teacher-student relationship. The questionnaire was administered in early spring. These questions were derived from the short version of the "Leerling Leerkracht Relatie Vragenlijst" (LLRV)

questionnaire of Koomen, Verschueren, and Pianta (2007). This Dutch questionnaire was based on the Student-Teacher Relationship Scale (STRS) of Pianta (2001). The questionnaire was assumed to contain three subscales: dependency, conflict, and closeness. The Dutch version of this questionnaire was validated with a sample of 2335 children aged 3 to 12. Overall, the 3-dimension model showed acceptable fit. Analysis of measurement invariance showed that generally scalar invariance did not hold, however, the revealed non-invariance for gender and age did not influence mean group comparisons (Koomen et al., 2012).

For the current dataset, a principal components factor analysis resulted in a three-factor solution, directly related to the three assumed subscales, explaining 73.6% of the variance (Driessen et al., 2009). The subscale dependency (5 items, $\alpha = .90$) related to the extent that the teacher perceived within the teacher-student relationship that the student was overly dependent on the teacher, for example, “This student asks for help from me even when it is not necessary”. The subscale closeness (5 items, $\alpha = .86$) related to the extent to which the teacher perceived the teacher-student relationship to reflect warmth, affection, and openness in communication. An example item was: “I share an affectionate, warm relationship with this child”. The subscale conflict (5 items, $\alpha = .93$) was related to the extent to which the teacher perceived the relationship as negative and adversarial, for example, “This child and I always seem to be struggling with each other”.

Descriptive statistics for the variables described above can be found in Table 1.

Analytic Strategy

Imputation of missing values. The original sample contained records of 9,881 students in 614 classes in 485 schools for whom the teachers' expectations were available. Of all values in the predictor variables that were of interest in the current study, 95.18% were complete, however, the missing values were distributed over many students. This resulted in

a pattern in which only 59.06% (5,836) of the total 9,881 students had a complete record for all the predictor variables of interest.

A comparison of complete cases compared to cases of students with one or more missing values yielded the following results. The teacher recommendations were slightly higher for the students with complete cases; $d = .05$, $t(8347) = 2.39$, $p = .016$. Similarly, the average score on the school leaver's test was slightly higher for the students with complete records; $d = .11$, $t(4029) = 3.36$, $p < .001$. Furthermore, the students with complete records performed higher on average on the mathematics test; $d = .15$, $t(8969) = 7.09$, $p < .001$, but not on the language; $d = .04$, $t(10059) = -1.76$, $p = .079$, and reading comprehension; $d = .03$, $t(10269) = 1.59$, $p = .113$ tests. With respect to the background characteristics, the complete record students were significantly different from students with missing values in terms of ethnic background; $\chi^2(2) = 24.12$, $p < .001$, and parental education; $\chi^2(2) = 17.24$, $p < .001$, but not for gender; $\chi^2(2) = 0.45$, $p = .832$. In all cases the differences between complete and incomplete cases were very small.

For the majority of students, very little information was missing. In such cases, the often-used method of listwise deletion of cases with missing values is wasteful of information and may lead to biased results (Graham, 2009). Therefore, we applied the multilevel multiple imputation Chained Equations technique (Van Buuren, 2011; Van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, 2006) to impute missing values using all available information. An imputation model was set up using the package MICE in R (Van Buuren & Groothuis-Oudshoorn, 2011). The imputation model was based on the following rules: 1) only predictor variables were imputed; the scores on track recommendations (dependent variable) were not imputed, 2) according to the measurement level of the variables appropriate methods were selected, logistic regression for the binary variables gender and ethnic background (logreg), polytomous regression (polyreg) for parental education, and predictive mean matching (pmm)

for the continuous variables such as student performance and dimensions of teacher-student relationships, 3) predictors were included in the imputation model if the bivariate correlation between the imputed variable and the predictor was larger than $r = 0.20$, and 4) predictors were included when there were at least 50% usable cases.

The distribution of the variables in the imputed datasets were very similar to the original distributions and convergence was achieved quickly for each of the imputed variables (i.e., within 10 iterations). We constructed five datasets with imputed values; this number was based on the percentage of missing values in the dataset. Results reported in the following tables are syntheses of five analyses run on these imputed data sets. For the parameter estimates and standard errors in the subsequent tables, the combination rules of Little and Rubin (2002) were used. Given that the differences between complete and incomplete cases were very small and that the between dataset variance (differences in the estimated coefficients of the predictor variables between the imputed datasets) was much smaller than the within dataset variance (standard errors of the coefficients) for all of the predictor variables, we expect only a very minor impact of the imputation of missing values on the results.

Multilevel modeling. The data had a hierarchical structure with students (level 1) nested within classes (level 2), and were analyzed using a two-level multilevel model (Snijders & Bosker, 2012), using the MLwiN 2.35 software (Rasbash, Charlton, Browne, Healy, & Cameron, 2009). School level was not included in the models because most schools in the sample had only one or two grade six classes, which was insufficient for an adequate distinction of the variance between the class and the school level.

To test whether teacher expectations and teacher reports of the dimensions of closeness, conflict, and dependency of the teacher-student relationships for older children (age 11-12) shared similar antecedents (H1), a multivariate multilevel regression model was

estimated. Four dependent variables were included in this analysis: track recommendations, perceived closeness, conflict, and dependency. First, an empty model was estimated to investigate the partitioning of variance over the student and the class level (Model 1). In the random part of the model, a full variance-covariance matrix was estimated at the student and the class level, allowing for the possible association between the four dependent variables at the student and class level. In a next step, student performance, background, and interactions between demographic variables and student performance were added to the model as predictor variables (Model 2). For continuous variables, such as the student performance variables, grand mean centering was applied. Separate coefficients were estimated for each of the dependent variables, allowing the coefficient of a predictor to vary between dependent variables in strength and direction. The coefficients of the demographic variables and the interactions from this model provided an indication of whether track recommendations and teacher-student relationship dimensions shared similar antecedents.

To test whether the teacher-student relationship was a precursor of track recommendations, a multilevel regression model was set up with teacher expectations as the dependent variable. Student performance, background and interactions between demographic variables and student performance (that were significant in the previous analysis, Model 2) were added to the model as covariates (Model 3). The dimensions of closeness, conflict, and dependency of the teacher-student relationships and their interactions with student performance and demographics were also added to the model. Despite the high correlation between student performance and track recommendations, it is important to include student performance as a covariate in the model because of substantial differences in cognitive skills between several subgroups of students. Not controlling for student performance would produce findings of bias towards (or against) gender, parental education, and minority status

across numerous student outcomes (Ferguson, 2003; Ready & Wright, 2011). Again, grand mean centering was applied for the continuous predictor variables.

In the random part of the latter model, we allowed a random intercept at class level as well as random slopes. By means of the random slopes, it is possible to investigate whether the coefficient varies among classes. First, we tested whether the random slopes of closeness, conflict, and dependency improved model fit on a one-by-one basis (Hox, Moerbeek, & van de Schoot, 2017). Compared to a model without random intercepts, a significant decrease in model misfit was observed when adding random slopes for conflict; $-2 \times \log \text{likelihood } 14736.06 - 14717.02, \chi^2(2) = 19.04, p < .001$, and for dependency; $-2 \times \log \text{likelihood } 14736.06 - 14721.20, \chi^2(2) = 14.86, p < .001$. Adding random slopes for closeness did not significantly improve model fit; $-2 \times \log \text{likelihood } 14736.06 - 14733.57, \chi^2(2) = 2.49, p = .288$. The final model (Model 3) therefore included random intercepts for the dependent variable track recommendations and random slopes at the class level for the main effects of conflict and dependency as well as their covariances. Deviance tests for this model indicated better fit compared to a no random slopes model ($-2 \times \log \text{likelihood } 14736.06 - 14709.18, \chi^2(5) = 28.88, p < .001$) and compared to the models with just random intercepts for conflict ($-2 \times \log \text{likelihood } 14717.02 - 14709.18, \chi^2(5) = 7.84, p = .049$) or for dependency ($-2 \times \log \text{likelihood } 14721.20 - 14709.18, \chi^2(5) = 12.02, p = .007$). There were no important changes in the coefficients of the fixed effects compared to the model without the random intercepts.

Additionally, differences between classes in intercepts and slopes are presented by means of 95% coverage intervals (Leckie, 2013). The derivation of 95% coverage intervals is based on the model assumption that the random effects are normally distributed, that is, the residuals of classes from the average (as estimated by the regression model) follow a normal distribution. Given normality, we expected 95% of the random effects for each level to lie in the range of ± 1.96 times the square root of the associated variance component. The variance

components referred to the class-level random intercept variance and class-level random slope variance from Model 3. As suggested by Leckie (2013), coverage intervals were presented centered round the corresponding coefficient from the fixed part of the model. It is important to note that coverage intervals do not provide information on the precision or significance of a model coefficient. The use of coverage intervals is merely a method to present differences between units at particular levels.

Results

Bivariate Relations

In Table 2, the zero-order correlations are presented between track recommendations, several measures of student performance, and the three dimensions of teacher-student relationships. The highest correlation was found between the teachers' track recommendations and the scores on the school leavers' test ($r = .89$). Somewhat lower, yet significant positive associations, were observed between the teachers' recommendations and the other test scores ($r = .53 - .74$). These high correlations indicated that the recommendations of the teachers were closely associated to the performance of the students during the final grade of primary education. Negative correlations were found between the dimensions of conflict and dependency and student performance (conflict $r = -.20$, dependency $r = -.30$) and track recommendations (conflict $r = -.20$, dependency $r = -.31$). The associations between closeness and student performance ($r = .06$) and track recommendations ($r = .07$) were positive but very close to zero.

Table 3 provides the descriptive statistics for the three dimensions of teacher-student relationships, track recommendations, and students' demographic background. For the teachers' track recommendation, the mean for the boys and girls was fairly similar; $t(62199) = 1.046, p = .296$. The differences in track recommendations between the groups based on parental education, however, were substantial; $F(2, 9878) = 925.410.18, p < .001$. The

average track recommendation of students in the group with high parental education was more than one track higher than that of the group of students with low parental education. Also, there appeared to be higher track recommendations for Dutch students compared to students from other ethnic backgrounds; $t(13647) = 18.120, p < .001$. However, in Table 3, student performance had not yet been taken into account; therefore, differences in track recommendations between the groups may have reflected differences in student performance.

When looking into differences between boys and girls for the three teacher-student relationship dimensions, teachers reported higher levels of closeness; $t(1363) = 18.130, p < .001$, and lower levels of conflict; $t(1141) = -16.640, p < .001$ with girls compared with boys. Differences between boys and girls with respect to dependency were not statistically significant $t(5886) = -1.913, p = .056$. Teachers reported lower levels of conflict; $F(2, 9878) = 57.591, p < .001$, and dependency; $F(2, 9878) = 84.283, p < .001$, the higher the level of parental education. Smaller though statistically significant differences were found for perceived closeness; $F(2, 8978) = 26.086, p < .001$. In the case of ethnic background, higher levels of perceived closeness; $t(288) = 8.862, p < .001$, and lower levels of perceived conflict; $t(724) = -7.726, p < .001$, and perceived dependency; $t(303) = -6.695, p < .001$, were reported by teachers for Dutch students compared to students from other ethnic backgrounds.

Student Background Variables as Antecedents of Track Recommendations and the Teacher-Student Relationship

This section provides the results regarding Hypothesis 1, and gives insight to what extent track recommendations and the teacher-student relationship shared similar antecedents. The results of the empty multivariate multilevel model are presented in Table 4 (Model 1). The variance components of the dependent variables can be found on the diagonal of the random part of Table 4, whereas the covariance between variance components of the dependent variables can be found beneath the variance components. For each of the

dependent variables, a substantial part of the variance was associated with the class level (track recommendations ICC = .129, closeness ICC = .203, conflict ICC = .212, dependency ICC = .270), indicating that classes differed in the average levels that teachers reported for all dependent variables. The relative differences between classes appeared the largest for the teachers' perception of dependency in the teacher-student relationship and the smallest for track recommendations.

The results of the multilevel model in which student performance and background characteristics were added as predictors are presented in Table 5 (Model 2). These variables explained 81.3% of the variance in track recommendations, 4.9% for closeness, 7.6% for conflict, and 5.9% for dependency, which implied that student performance and the demographic background were far stronger predictors of track recommendations compared to the teacher-student relationships.

Student performance. Student performance seemed a consistent predictor of teachers' track recommendations and the dimensions conflict and dependency, whereby track recommendations and the perception of conflict were significantly predicted by all four of the performance variables and dependency was predicted by three of the four performance variables. This indicated that teachers gave higher recommendations, and had lower perceived conflict and dependency for higher performing students. With respect to the last teacher-student relationship dimension closeness, there seemed to be no clear pattern in the associations with student performance. Of the four variables, one significant result appeared; mathematics test $b = .002$, $\beta = 0.027$, $t(9788) = 2.00$, $p = .045$. For all performance variables, the coefficients were very close to zero.

Gender. With respect to gender, only significant associations were present for closeness and conflict. On average, teachers reported higher levels of closeness towards girls compared to boys; $b = .212$, $\beta = 0.321$, $t(9788) = 10.10$, $p < .001$. For students with average

performance, the reported conflict for girls was considerably lower than for boys; $b = -.259$, $\beta = 0.320$, $t(9788) = 10.36$, $p < .001$. However, regarding conflict, the interaction between gender and the school leavers' test indicated that the association between conflict and student performance showed a stronger negative association for boys than for girls; $b = .008$, $\beta = 0.099$, $t(9788) = 4.00$, $p < .001$.

Ethnic Background. Ethnicity was a significant predictor of track recommendations and closeness, be it in a different direction. On average, teachers gave higher recommendations to minority students; $b = .056$, $\beta = 0.046$, $t(9788) = 2.00$, $p = .045$, compared to Dutch students. The differences in track recommendations between Dutch and minority students were larger for high performing students in comparison to low performing students; $b = .003$, $\beta = 0.025$, $t(9788) = 3.00$, $p = .003$. Additionally, teachers reported lower levels of closeness towards minority students compared to students with Dutch backgrounds; $b = -.156$, $\beta = 0.241$, $t(9788) = 4.59$, $p < .001$. The difference between minority and Dutch students in reported closeness by teachers was smaller when the students' performance was higher; $b = .006$, $\beta = 0.091$, $t(9788) = 3.00$, $p = .003$. Although for the average performing students, no significant differences in perceived conflict were found between Dutch and minority students, ethnicity was involved in ones significant interaction. The negative association between performance and conflict seemed stronger for minority students compared to Dutch students; $b = -.005$, $\beta = -0.06$, $t(9788) = 2.50$, $p = .012$,

Parental education. Concerning parental education, significant associations were found with teachers' track recommendations, closeness, and conflict, all associations in the expected directions. Teachers reported higher recommendations for students from highly educated parents; $b = .086$, $\beta = 0.071$, $t(9788) = 4.09$, $p < .001$, and lower expectations for students from less educated parents; $b = -.126$, $\beta = -0.104$, $t(9788) = 5.25$, $p < .001$, compared to the reference group (middle SES). However, it seemed that the association between

performance and track recommendations was stronger for the reference group than for students from less educated parents; $b = -.006$, $\beta = -0.049$, $t(9788) = 6.00$, $p < .001$, or students from highly educated parents; $b = -.002$, $\beta = -0.017$, $t(9788) = 2.00$, $p = .045$. Regarding closeness, teachers reported lower levels of closeness for students of less educated parents compared to the reference group (middle); $b = -.076$, $\beta = 0.115$, $t(9788) = 2.92$, $p = .004$. Finally, higher levels of perceived conflict were found for students from less educated parents compared to the reference group; $b = .096$, $\beta = 0.119$, $t(9788) = 3.00$, $p = .003$.

Association Between Teacher-Student Relationships and Track Recommendations

This section provides the results regarding Hypothesis 2. In Table 6, the results of a multilevel regression model are presented, in which the three dimensions of the teacher-student relationship were used as additional predictors of teachers' track recommendations (Model 3), after taking into account the demographic background variables that showed a significant association with teacher expectation (see Model 2). The teacher-student relationship as perceived by the teacher added very little to the explanation of track recommendations ($\Delta R^2 = .005$). Compared to the previous model for the dependent variable track recommendations (Table 5), the results for student performance and student background did not change.

For closeness and conflict, no statistically significant main and interaction effects were found. For dependency, both a main effect and an interaction effect with student performance were found to be statistically significant; main effect $b = -.081$, $\beta = -0.054$, $t(5806) = 4.76$, $p < .001$; interaction school leavers' test $b = -.003$, $\beta = -0.021$, $t(5806) = 3.00$, $p = .001$. These findings implied that for the average student, a weak negative association was found between the teachers' perceptions of dependency and track recommendations, however, the strength of the association was dependent on the students' performance. Teacher-reported dependency

had a slightly stronger negative association with teacher expectations for high performing students compared to low performing students.

In Table 7, the results are presented with respect to the random intercepts and slopes from Model 3 by means of 95% coverage intervals (see Leckie, 2013). With respect to the random intercepts, assuming a normal distribution of between-class differences, 95% of the classes are expected to lie in the range of track recommendations between 3.721 and 4.579¹. Thus, in classes at the 97.5th percentile of the class distribution, students from the reference groups receive track recommendations 0.858 points higher than similar students in classes at the 2.5th percentile. This implied that the differences between high and low expectation teachers added up to almost one track. Regarding conflict, 95% of the classes were expected to lie in the range -0.098 to 0.078. This implied that in some classes conflict was weighted positively in the track recommendations whereas in other classes it was weighted negatively. The variation between classes for conflict was modest compared to variation between classes for dependency. For dependency, 95% of the classes were expected to lie in the range -0.189 to 0.027² which indicated that, dependency was weighted negatively in the track recommendation for the vast majority of the classes.

Discussion

In this study, we aimed to investigate (1) whether teacher expectations and teacher reports of the dimensions of closeness, conflict, and dependency of the perceived teacher-student relationships with older children (aged 11-12) shared similar antecedents, and (2) whether the dimensions of closeness, conflict, and dependency of the perceived teacher-

¹ This 95% coverage interval was derived from the intercept (4.150, Table 5) and the intercept variance (0.048, Table 6) from Model 3. The lower bound of the coverage intervals is calculated as $4.150 - 1.960 * \sqrt{0.048} = 3.271$, whereas the upper bound is $4.150 + 1.960 * \sqrt{0.048} = 4.579$.

² It is important to note that coverage intervals do not provide information on the precision or significance of a model coefficient, it is merely a method to present differences between units at particular levels. The fact that the 95% coverage interval of dependency does include zero, does therefor not mean that dependency is not significant as predictor of track recommendations.

student relationship predicted teacher expectations, after the performance of students was taken into account. Therefore, we analyzed data from a large sample of 9,881 students in the final grade of Dutch primary education, using track recommendations as a special case of teacher expectations.

Based on previous research findings with respect to teacher expectations and teacher-student relationships, we expected gender, parental education, and ethnic background to be antecedents (H1). This hypothesis that expectation and the teacher-student relationship shared antecedents was not confirmed. Although several significant relations were found between the suggested antecedents and track recommendations and the dimensions of teacher-student relationships, the pattern of associations was not consistent showing more difference than similarities. For example, none of these suggested antecedents were significantly related to the dimension dependency, whereas they showed relations to the other outcomes.

For gender, the relationships were in the expected directions with higher perceived closeness for girls and higher perceived levels of conflict for boys, but no association was present for track recommendations. These findings corresponded to previous teacher-student relationship research based on samples of young children (Birch & Ladd, 1998, Silver et al., 2005). With respect to the students' parental education, the associations were in the expected direction and consistent over the outcome variables with the exception of dependency. The differences between the low and middle group of parental education appeared statistically significant for the teacher-student relationship dimensions, indicating higher levels of conflict and lower levels of closeness and track recommendations for students whose parents were lowly educated compared to the middle education group. Additionally, differences were found between the middle and high group of parental education for track recommendations.

Regarding the students' ethnic background, contrasting findings were found for track recommendations compared to the teacher-student relationship. Although teachers in this

sample characterized their relationship with minority students by lower levels of closeness, their recommendations for minority students were, on average, higher than for majority students, all other characteristics being equal. In the international context, minority status is usually related to lower teacher expectations (e.g., Glock & Krolak-Schwerdt, 2013; McKown & Weinstein, 2008; Rubie-Davies, Hattie, & Hamilton, 2006; Tenenbaum & Ruck, 2007), however, in the Netherlands, track recommendations for minority students have systematically appeared higher than for Dutch majority students (e.g., Inspectorate of Education, 2014; Timmermans et al., 2015). Thus far, there is no conclusive explanation for this finding. Higher track recommendations for minority students may be a consequence of positive discrimination (De Jong, 1987), a fear of teachers being accused of racism (Jungbluth, 1985; Stevens, 2008).

The results concerning the relationship between the perceived teacher-student relationship and student academic performance seemed not always consistent with previous findings. In this study, there seemed to be no clear pattern in the relations between perceived closeness and student performance, whereas previous research including in the Dutch context (e.g., Jellesma, Zee, & Koomen, 2015; Ryan, Stiller, & Lynch, 1994; Wentzel, 1997) has shown that higher levels of closeness coincided with more favorable outcomes for both younger and older students. Corresponding to previous findings, higher levels of reported conflict coincided with lower levels of student performance. Furthermore, although the role of dependency for older students has not yet been fully established (Ang, 2005), we found a consistent pattern between student performance and dependency, whereby lower levels of student performance were associated with higher levels of reported dependency. From the teacher's perspective, relationships characterized by negativity (conflict and overdependency) may lead to frequent attempts to control students' behavior and thus hinder efforts to promote a positive school environment for them (Hamre & Pianta, 2001). Additionally, the quality of

the relationships may reflect the extent to which children are able to engage with the instructional resources present in classrooms (Entwisle & Hayduk, 1988). Both may be explanations of the negative association between students' academic performance and the negative dimensions of the teacher-student relationship. On the other hand, students' academic performance appeared a very strong predictor of track recommendations, which is in line with both research on track recommendations (e.g., Timmermans et al., 2016, Timmermans, de Boer, Amsing, & Van der Werf, 2018) as well as research on teacher expectations in general (e.g., Jussim & Harber, 2005).

For our second hypothesis, we expected the perceived teacher-student relationship to be associated with teacher expectations (H2), because it could be the teacher's perceived relationship that is principally influential in educational decisions, the marks that are given, and actions toward students in instructive and interpersonal interactions (Ang, 2005; Hamre & Pianta, 2001), culminating, in the end, with higher recommendations. The results indicated that teachers' track recommendations were only dependent to a very limited degree on the perceived quality of the interpersonal dyadic relationship between teacher and student ($\Delta R^2 = .005$). Given the importance of track recommendations for students' future school careers, it is a reassuring finding that teachers strongly rely on objective data on students' academic performance, rather than on their subjective perceptions on how well they like the students.

Even though we expected a generally positive relationship between perceived closeness and teacher expectations (based on Hughes et al., 2005; Rubie-Davies, 2010; Timmermans et al., 2016), we did not find main or interaction effects of teachers' perceived closeness with student background. A possible explanation is the time of assessment of the expectation in the current study. The track recommendations provided by the teachers were measured late spring, thus almost 8 months after the start of the school year. If one assumes that teachers continuously adjust their expectations to align them with student achievement

(Brophy, 1983), there may be less room for other predictors than performance to be related to expectations the further the measurement extends into the school year. No previous research was available – to our knowledge – for the association between expectations and the two negative relational factors: conflict and dependency. Nonetheless, the expected negative association was only found for perceived dependency. After the performance of the students was taken into account; teachers had lower expectations for those students with whom they perceived a relationship characterized by higher levels of dependency. However, this negative association was stronger for high performing than for low performing students. A possible explanation is that teachers expect high achieving students to be more independent. If high performing students are displaying high levels of dependency and seek a lot of support from their teachers, the teachers may perceive these students as less capable. Another interpretation could be that teachers assume that a certain level of independence is necessary to succeed in secondary education, especially in the higher tracks.

From the resulting differences between classes, it was evident that the recommendations given in some classes were higher than in other classes, almost reaching one track between classes at the 2.5th and 97.5th percentile. Furthermore, there appeared some variation between classes in the coefficients of conflict, and dependency, indicating that the weighting of these variables differed from teacher to teacher. Conflict was weighted positively in some classes whereas, in other classes, they were weighted negatively. Dependency, however, was weighted negatively in the track recommendation for the vast majority of classes. The differences between classes were similar to those found in previous research (Timmermans et al., 2015; 2016), and more substantial than the average coefficients from the sample. Especially the between class differences warrant future research. First, the random slopes of conflict and dependency indicate that the track recommendations are based on different sources or a different weighting of the sources when formulating track

recommendation. Second, the between class differences in track recommendations imply that for children with comparable performance teachers may give very different recommendations, which has a major influence on the future academic careers. If the interpersonal teacher-student relationship has only a minor impact on track recommendations, future research should search which characteristics of teacher or the context contributes to these differences in track recommendation and educational opportunities.

Limitations

In interpreting the results of this study, a number of limitations need to be considered. The nature of this large-scale dataset does not permit any conclusions regarding causal pathways between the teachers' perceptions of the teacher-student relationship and teacher expectations. In our study, we focused on the prediction of teacher expectations based on the teacher-student relationship, as indicated in suggestions and indications from previous research (Ang, 2005; Hamre & Pianta, 2001). We can be confident of the correct temporal precedence in this study, due to the difference stages of data collection.

The current study has filled a gap in the teacher expectation and teacher-student relationship literature in which the interpersonal character of the relationship has been largely neglected. However, it is important to consider that a number of potential predictors could not be included in this study. Antecedents not tested in this study include characteristics of the teachers (teacher's experience, background and beliefs) and other perceptions of student attributes (perceived student behavior, students' physical appearance). This study is therefore only a step in the direction of understanding contributing factors, but it is important to develop an overarching framework of factors contributing to teacher expectations and teacher-student relationships. Future research may benefit from studies establishing how teacher expectations and the teacher-student relationship develop over the course of a school year (e.g., Kuklinski & Weinstein, 2000; Martinek, 1980), which comprehensive set of factors

predict expectations and the teacher-student relationship, and whether the influence of contributing factors is dependent on the time of year. For example, it seems plausible that factors like student demographics may be more predictive of expectations at the beginning of the year, and the relationship less so, given that a relationship has yet to be established. This requires a longitudinal assessment of teacher expectations, student performance, and teacher perceptions on student behavior as well as the perceived relationship.

Information on teachers' track recommendations and their perceptions of the teacher-student relationship were both based on teacher reports. As a result, the method of measuring both constructs may have accentuated the relations among these variables. Another limitation also stems from the measurement method because the teacher reports did not allow us to distinguish between the perceptions of teachers and the actual teacher-student relationship. Although it is important to note that it is the perception of the teacher that is crucial in shaping expectations and not the relationship itself, we cannot be certain that the perceived and actual relationships matched.

Conclusion

To conclude, this study is one of the first papers that has actually addressed the interpersonal character of teacher expectations and the findings indicate that teachers strongly rely on objective data on students' academic performance in forming track recommendations and that their subjective perceptions on how well they like the students only played a minor role in the negative association with the dimension dependency. However, this study also indicates that a different weighting of the conflict and dependency applied between teachers and classes when formulating track recommendation, indicating that it did play a (stronger) role in some of the classes. The results strengthen the hypothesis that the perceived dyadic teacher-student relationship is to some extent a precursor of teacher expectations. Following these findings, future research concerning teacher expectation accuracy would benefit from

investigating both teacher and student characteristics as well as their dyadic relationship.

Future research might also further investigate the possible mediating effect of teacher expectations in the associations between the teacher-student relationship and various student outcomes.

References

- Aarsen, E. van, Roeleveld, J., & Luyten, J. W. (2013). *Voorspellende waarde van LOVS toetsen voor schoolsucces. [The predictive value of Monitoring and Evaluation System tests for school success]*. Utrecht: Oberon.
- Alvridez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology*, 91, 731–746.
- Ang, R. P. (2005). Development and validation of the Teacher–Student Relationship Inventory using exploratory and confirmatory factor analysis. *The Journal of Experimental Education*, 74, 55–73.
- Babad, E. (1992). Teacher expectancies and nonverbal behavior. In: S. Feldman, (Ed.), *Applications of nonverbal behavioral theories and research* (pp. 167-190). Hillsdale, NJ: Erlbaum.
- Babad, E. Y. (2009). *The social psychology of the classroom*. New York: Routledge.
- Birch, S. H., & Ladd, G. W. (1997). The teacher–child relationship and children's early school adjustment. *Journal of School Psychology*, 35, 61–79.
- Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher–child relationship. *Developmental Psychology*, 34, 934–946.
- Bishop, R. & Berryman, M. (2006). *Culture speaks: Cultural relationships and classroom learning*. Wellington, New Zealand: Huia.
- Bonvin, P., & Genoud, P. A. (2006, September). *Teacher perceptions of student characteristics: “Halo effects” between behavioural and social perceptions, and expectations for academic achievement*. Paper presented at European Conference on Educational Research (ECER), Geneva, Switzerland.
- Bosker, R.J. & Van der Velden, R.K.W. (1985) Schooleffecten en rendementen (Groningen (RION, Instituut voor Onderwijsonderzoek Rijksuniversiteit Groningen)).

- Bosker, R.J. & Van der Velden, R.K.W. (1989) The effects of secondary schools on the educational careers of disadvantaged pupils (RION, Institute for Educational Research, Groningen).
- Boxtel, H.W. van, Engelen, R. & Wijs, A. de (2011). *Wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010. [Scientific report on the School Leavers test 2010]*. Arnhem, The Netherlands: Cito.
- 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.
- Brophy, J. E., & Good, T. L. (1974). *Teacher-student relationships: Causes and consequences*. New York, NY: Holt, Rinehart, & Winston.
- Collins, W. A. & Repinski, D. J. (1994). Relationships during adolescence: Continuity and change in interpersonal perspective. In R. Montemayor, G. Adams, & T. Gullotta (Eds.), *Personal relationships during adolescence* (pp. 7-36). Thousand Oaks, CA: Sage.
- 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
- Doumen, S., Verschueren, K., Buyse, E., Germeijs, V., Luyckx, K., & Soenens, B. (2008). Reciprocal relations between teacher–child conflict and aggressive behavior in kindergarten: A three-wave longitudinal study. *Journal of Clinical Child and Adolescent Psychology*, 37, 588–599.
- Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J., & van der Veen, I. (2009). *Cohortonderzoek COOL⁵⁻¹⁸. Technisch rapport basisonderwijs, eerste meting 2007/08*. Nijmegen / Amsterdam, The Netherlands: ITS / SCO-Kohnstamm Instituut

- Entwisle, D. R. & Alexander, K. L. (1988). Factors affecting achievement test scores and marks of black and white first graders. *Elementary School Journal*, 88, 449–471.
- Entwisle, D. R., & Hayduk, L. A. (1988). Lasting effects of elementary school. *Sociology of Education*, 61, 147–159.
- Ferguson, R. F. (2003). Teachers' perceptions and expectations and the black-white test score gap. *Urban Education*, 38, 460 – 507.
- Feron, E., Schils, T., & ter Weel, B. (2016). Does the teacher beat the test? The value of the teacher's assessment in predicting student ability. *De Economist*, 164, 391–418. doi: 10.1007/s10645-016-9278-z
- 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.
- Graham, J. W. (2009). Missing data analysis: Making it in the real world. *Annual Review of Psychology*, 60, 549 – 576.
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, 625–638.
- Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, 97(3), 363-386.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
- Hinde, R. A. (1997). *Towards understanding relationships*. London, UK: Academic Press.
- 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.

- Hoge, R. B. (1984). The definition and measurement of teacher expectations: Problems and prospects. *Canadian Journal of Education*, 9, 213–228.
- Hox, J. J., Moerbeek, M. & van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications, Third Edition*. New York: Routledge.
- Hughes, J. N., Cavell, T. A., & Jackson, T. (1999). Influence of student–teacher relationship on childhood aggression: A prospective study. *Journal of Clinical Child Psychology*, 28, 173–184.
- Hughes, J. N., Gleason, K. A., & Zhang, D. (2005). Relationship influences on teachers' perceptions of academic competence in academically at-risk minority and majority first grade students. *Journal of School Psychology*, 43, 303–320.
- Inspectorate of Education (2007). *Onderadvisering in beeld. [Teacher recommendations in the picture]*. Utrecht, The Netherlands: Author.
- Inspectorate of Education (2014). *De kwaliteit van het basisschooladvies. [The quality of track recommendations]*. Utrecht, The Netherlands: Author.
- Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS Rekenen-Wiskunde voor groep 3 tot en met 8. [Scientific validation of the mathematics tests from the monitoring system grade 1 until grade 6]*. Arnhem, The Netherlands: Cito.
- Jellesma, F. C., Zee, M., & Koomen, H. M. Y. (2015). Children's perceptions of the relationship with the teacher: Associations with appraisals and internalizing problems in middle childhood. *Journal of Applied Developmental Psychology*, 36, 30–38.
- Jong, M. J. de (1987). *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis- naar voortgezet onderwijs. [Origins, knowledge and Chances. Immigrant and Dutch students during the transition from primary to secondary education]*. Lisse, The Netherlands: Swets & Zeitlinger B.V.

- Jungbluth, P. (1985) *Verborgen discriminatie: Leerlingbeeld en onderwijsaanbod op de basisschool [Hidden discrimination: Views on students and educational provisions in primary education]*. Nijmegen, The Netherlands: ITS.
- Jussim, L. & Eccles, J. S. (1992) Teacher expectations II: Construction and reflection of student achievement. *Journal of Personality and Social Psychology*, 63, 947–961.
- Jussim, L., Eccles, J. S., & 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*, 29, 281–388.
- 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.
- Koepke, M. F. & Harkins, D. A. (2008). Conflict in the classroom: Gender differences in the teacher–child relationship. *Early Education and Development*, 19, 843 – 864.
- Koomen, H.M.Y., Verschueren, K., & Pianta, R.C. (2007) *Leerling leerkracht relatie vragenlijst. Handleiding. [Student teacher relationship questionnaire. Manual]*. Houten, The Netherlands: Bohn Stafleu van Loghum.
- Koomen, H. M. Y., Verschueren, K., van Schooten, E., Jak, S., & Pianta, R.C. (2012). Validating the Student-Teacher Relationship Scale: Testing factor structure and measurement invariance across child gender and age in a Dutch sample. *Journal of School Psychology*, 50, 215–234.
- Leckie, G. (2013). *Three-Level Multilevel Models: Concepts. LEMMA VLE Module 11*. Bristol, UK: Centre for Multilevel Modelling, University of Bristol.
- Laible, D. & Thompson, R. A. (2007). Early socialization: A relationship perspective. In J. Grusec & P. Hastings (Eds.), *Handbook of social development* (pp. 181-207). New York, NY: Guilford.

- Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, 72, 791–809.
- Madon, S., Willard, J., Guyll, M., & Scherr, K. C. (2011). Self-fulfilling prophecies: Mechanisms, power and links to social problems. *Social and Personality Psychology Compass*, 5, 578–590.
- Mantzicopoulos, P., & Neuharth Pritchett, S. (2003). Development and validation of a measure to assess head start children's appraisals of teacher support. *Journal of School Psychology*, 41, 431–451.
- McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student-teacher relationships and their outcomes. *Educational Research Review*, 14, 1 – 17. doi:10.1016/j.edurev.2014.12.001
- McKnown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, 46, 235–261.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8, 193–210.
- Ministry of Education (2014). *Key-figures 2009-2013: Education, culture, science*. Den Haag, The Netherlands: Author.
- Nurmi, J. E. (2012). Students' characteristics and teacher-child relationships in instruction: A meta-analysis. *Educational Research Review*, 7, 177 – 197. doi:10.1016/j.edurev.2012.03.001
- OECD (2016). *Netherlands 2016: Foundations for the future. Reviews of national policies for education*. Paris: OECD. doi: 10.1787/9789264257658-en
- Palermo, F., Hanish, L. D., Martin, C. L., Fabes, R. A., & Reiser, M. (2007). Preschoolers' academic readiness: What role does the teacher– child relationship play? *Early Childhood Research Quarterly*, 22, 407–422.

- Pianta, R. C. (2001). *Student–teacher relationship scale: Professional manual*. Psychological Assessment Resources, Lutz, FL
- Pianta, R. C., & Steinberg, M. (1992). Teacher–child relationships and the process of adjusting to school. In R. C. Pianta (Ed.), *Beyond the parent: The role of other adults in children's lives. New directions for child development* (pp. 61-79). San-Francisco, CA: Jossey-Bass.
- Pianta, R. C., Steinberg, M. S., & Rollins, K. B. (1995). The first two years of school: Teacher–child relationships and deflections in children's classroom adjustment. *Development and Psychopathology*, 7, 298–312.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M. & Cameron, B. (2009). *MLwiN Version 2.1*. Bristol, UK: 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.
- 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) Doelgroeppeerlingen in het basisonderwijs. Historische ontwikkeling en actuele situatie. [At risk students in primary education] (Amsterdam, Kohnstamm Instituut).
- Rubie-Davies, C. M. (2008). Teacher expectations. In T. Good (Ed.), *21st century education: A reference handbook* (pp. I-254 – I-265). Thousand Oaks: SAGE Publications. doi: 10.4135/9781412964012.n27

- Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology*, 80, 121–135. 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
- Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents and friends as predictors of academic motivation and self-esteem. *Journal of Early Adolescence*, 14, 226–249.
- Saft, E. W. & Pianta, R. C. (2001). Teachers' perceptions of their relationships with student: Effects of child age, gender, and ethnicity of teachers and children. *School Psychology Quarterly*, 16, 125–141.
- Schilt-Mol, T. M. M. L. van (2007). *Differential Item Functioning en itembias in de Cito-Eindtoets Basisonderwijs*. Amsterdam, The Netherlands: Aksant.
- [dataset] SCO-Kohnstamm Instituut Amsterdam; ITS Radboud Universiteit Nijmegen (2007): *Cohortonderzoek Onderwijsloopbanen van 5-18 jaar - COOL 5-18 - Basisonderwijs 2007/08*. DANS.<https://doi.org/10.17026/dans-xxs-4h2a>
- Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and teacher–child relationship during the school transition. *Journal of School Psychology*, 43, 39–60.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London/Thousand Oaks/New Delhi: Sage Publications.

- Stevens, P. A. J. (2008). Exploring pupils' perceptions of teacher racism in their context: a case study of Turkish and Belgian vocational education pupils in a Belgian school, *British Journal of Sociology of Education*, 29, 175–187.
- Stevens, P. A. J., Clycq, N., Timmerman, C., & Houtte, M. van (2011). Researching race/ethnicity inequality in the Netherlands: a critical review of the research literature between 1980 and 2008. *British Educational Research Journal*, 37, 5 – 43. doi: 10.1080/01411920903342053
- 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.
- Thijs, J. T., Koomen, H. M. Y., & van der Leij, A. (2008). Teacher–child relationships and pedagogical practices: Considering the teacher's perspective. *School Psychology Review*, 37, 244–260.
- Thijs, J. T., Westerhof, S., & Koomen, H. M. Y. (2012). Ethnic incongruence and the student–teacher relationship: The perspective of ethnic majority teachers. *Journal of School Psychology*, 50, 257 – 273.
- Timmermans, A. C., Kuyper, H., & Werf, M. P. C. van der (2013). *Schooladviezen en onderwijsloopbanen: Voorkomen, risicofactoren en gevolgen van onder- en overadvisering*. Groningen: GION.
- Timmermans, A. C., Kuyper, H., Van der Werf, M. P. C. (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, 459–478. doi: 10.1111/bjep.12087
- Timmermans, A. C., De Boer, H., & Van der Werf, M. P. C. (2016). An investigation of the relationship between teachers' expectations and teachers' perceptions of student

- attributes. *Social Psychology of Education*, 19, 217–240. doi: 10.1007/s11218-015-9326-6
- Timmermans, A. C., de Boer, H., Amsing, H. T. A., & van der Werf, M. P. C. (2018). Track recommendation bias: Gender, migration background and SES bias over a 20-year period in the Dutch context. *British Educational Research Journal*, 44(5), 847-874. <https://doi.org/10.1002/berj.3470>
- van Berkel, S., Hilte, M., Groenen, I., & Engelen, R. (2014). *Wetenschappelijke verantwoording Woordenschat groep 7 en 8. [Scientific validation of the vocabulary tests grade 5 and grade 6]*. Arnhem, The Netherlands: Cito.
- van Boxtel, H., Engelen, R., & de Wijs, A. (2011). *Wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010. [Scientific validation of the school leavers test 2010]*. Arnhem, The Netherlands: Cito
- Van Buuren, S. (2011). Multiple Imputation of Multilevel Data. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 173-196). New York, NY: Routledge.
- Van Buuren, S. & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-67.
- Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification of multivariate imputation. *Journal of Statistical Computation and Simulation*, 76, 1049 – 1064.
- Verhelst, N.D., & Glas, C.A.W. (1995). The one parameter logistic model. In: G.H. Fischer & I.W. Molenaar (Eds.), *Rasch models: Foundations, recent developments and applications* (pp. 215-239). New York, NY: Springer.
- Verhelst, N. D., Glas, C. A. W. & Verstralen, H. H. F. M. (1995). *OPLM: One parameter logistic model. Computer program and manual*. Arnhem: Cito.

- Weekers, A., Groenen, I., Kleintjes, F., & Feenstra, H. (2011). *Wetenschappelijke verantwoording papieren toetsen Begrijpend lezen voor groep 7 en 8. [Scientific validation of the paper tests for reading comprehension grade 5 and grade 6]*. Arnhem, The Netherlands: Cito.
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.
- Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, 89, 411–419.
- Wentzel, K. R. (2012). Teacher-student relationships and adolescent competence at school. In T. Wubbels, P. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 19-36). Rotterdam, The Netherlands: Sense Publishers.
- Zee, M. & Koomen, H. M. Y. (2017). Similarities and dissimilarities between teachers' and students' relationship views in upper elementary school: The role of personal teacher and student attributes. *Journal of School Psychology*, 64, 43 – 60.

Table 1

Descriptive Statistics of Variables used in the Analysis

Student level variables		Mean	SD	Min.	Max.	%
Teacher expectations		4.20	1.21	1.50	6.00	
School leavers test		533.13	10.06	501.00	550.00	
Language test		111.71	14.19	9.00	177.00	
Mathematics test		116.92	8.93	54.00	160.00	
Reading comprehension test		54.79	15.95	1.00	100.00	
Closeness		3.48	0.66	1.00	5.00	
Conflict		1.77	0.81	1.00	5.00	
Dependency		2.08	0.77	1.00	5.00	
Gender	Girls					49.4
	Boys					50.6
Parental education	High					25.5
	Middle					41.0
	Low					33.5
Ethnic background	Dutch majority					73.6
	Minority					26.4

Note. $N = 5836$ based on complete cases.

Table 2

Bivariate Correlations

	1	2	3	4	5	6	7
1 Teacher expectations							
2 School leavers test	.89*						
3 Language test	.53*	.54*					
4 Mathematics test	.73*	.78*	.40*				
5 Reading comprehensions test	.74*	.76*	.58*	.59*			
6 Closeness	.07*	.06*	.07*	.04*	.08*		
7 Conflict	-.20*	-.20*	-.12*	-.15*	-.20*	-.35*	
8 Dependency	-.31*	-.30*	-.18*	-.25*	-.26*	.01	.47*

$N = 9881$; * $p < .001$

Table 3

Between Group Differences in Teacher-Student Relationship and Teacher Expectations

Subgroups		Teacher expectations	Closeness	Conflict	Dependency
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Gender	Girls	4.16 (1.25)	3.61 (0.68)	1.63 (0.74)	2.07 (0.79)
	Boys	4.18 (1.24)	3.36 (0.67)	1.92 (0.89)	2.10 (0.81)
Parental education	High	4.85 (1.06)	3.50 (0.67)	1.68 (0.78)	1.95 (0.76)
	Middle	4.20 (1.17)	3.52 (0.67)	1.74 (0.81)	2.06 (0.79)
	Low	3.55 (1.17)	3.42 (0.72)	1.90 (0.89)	2.22 (0.82)
Ethnic background	Dutch	4.30 (1.22)	3.52 (0.67)	1.74 (0.80)	2.05 (0.79)
	Minority	3.77 (1.24)	3.37 (0.73)	1.89 (0.90)	2.18 (0.81)

N = 9881

Table 4

Results of the Empty Multivariate Multilevel Regression Model (Model 1)

Response	Teacher expectations		Closeness		Conflict		Dependency	
	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>
Fixed Part								
Intercept	4.153*	0.022	3.484*	0.014	1.786*	0.017	2.096*	0.018
Random Part								
<i>Class level</i>								
Teacher expectations	0.200	0.017						
Closeness	-0.009	0.008	0.095	0.007				
Conflict	-0.040	0.010	-0.047	0.007	0.146	0.011		
Dependency	-0.033	0.010	-0.026	0.006	0.115	0.009	0.171	0.012
<i>Student level</i>								
Teacher expectation	1.354	0.020						
Closeness	0.067	0.007	0.373	0.006				
Conflict	-0.163	0.009	-0.150	0.005	0.542	0.008		
Depencency	-0.271	0.009	0.027	0.004	0.191	0.006	0.462	0.007
Model fit								
-2*loglikelihood	52,088.420							

Note. Number of classes = 614. Number of students = 9881. Number of responses = 39524. * $p < .05$

Table 5

Results of the Multivariate Multilevel Regression Model with the Predictors (Model 2)

Response	Teacher expectations		Closeness			Conflict		Dependency	
	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>	<i>B</i>	<i>SEb</i>	
Fixed Part									
Intercept	4.156*	0.016	3.436*	0.021	1.878*	0.024	2.113*	0.024	
School leavers test	0.083*	0.002	-0.002	0.002	-0.011*	0.002	-0.017*	0.002	
Language test	0.005*	0.001	0.001	0.001	0.002*	0.001	0.000	0.001	
Mathematics test	0.016*	0.001	0.002*	0.001	-0.004*	0.002	-0.006*	0.001	
Reading comprehension test	0.008*	0.001	0.002	0.001	-0.003*	0.001	-0.002*	0.001	
Gender (girl)	0.021	0.018	0.212*	0.021	-0.259*	0.025	-0.040	0.023	
Ethnicity (minority)	0.056*	0.028	-0.156*	0.034	0.052	0.042	0.024	0.036	
Parental education (Low)	-0.126*	0.024	-0.076*	0.026	0.096*	0.033	0.031	0.029	
Parental education (High)	0.086*	0.021	-0.023	0.026	0.036	0.030	0.008	0.027	
Gender (girl) * Ethnicity (minority)	-0.007	0.026	0.087	0.031	-0.048	0.042	-0.054	0.035	
Gender (girl) * Parental education (Low)	0.001	0.028	0.041	0.035	-0.019	0.041	0.025	0.037	
Gender (girl) * Parental education (High)	0.012	0.027	0.002	0.033	-0.021	0.040	0.029	0.037	
Gender (girl) * School leavers test	-0.000	0.002	-0.000	0.001	0.008*	0.002	0.002	0.001	
Ethnicity (minority) * School leavers test	0.003*	0.001	0.006*	0.002	-0.005*	0.002	0.003	0.002	
Parental education (Low) * School leavers test	-0.006*	0.001	0.000	0.002	0.001	0.002	0.002	0.002	
Parental education (High) * School leavers test	-0.002*	0.001	0.002	0.002	-0.003	0.002	-0.003	0.002	

Ethnicity (minority) * Parental education (Low)	0.024	0.030	0.007	0.037	-0.052	0.045	-0.002	0.042
Ethnicity (minority) * Parental education (High)	-0.066	0.040	-0.025	0.048	0.018	0.057	0.038	0.053
Random Part								
<i>Class level</i>	σ^2	$SE\sigma^2$	σ^2	$SE\sigma^2$	σ^2	$SE\sigma^2$	σ^2	$SE\sigma^2$
Teacher expectations	0.049	0.004						
Closeness	-0.002	0.004	0.095	0.007				
Conflict	0.002	0.004	-0.044	0.006	0.136	0.010		
Dependency	-0.002	0.005	-0.025	0.006	0.105	0.009	0.166	0.001
<i>Student level</i>								
Teacher expectation	0.242	0.004						
Closeness	0.008	0.003	0.350	0.005				
Conflict	-0.017	0.005	-0.124	0.005	0.500	0.008		
Depencyency	-0.028	0.004	0.040	0.005	0.157	0.005	0.407	0.006
Model fit								
-2*loglikelihood	41,610.552							

Note. Number of classes = 614. Number of students = 9,881. Number of responses = 39,524. * $p < .05$

Table 6

Results of the Multilevel Regression Model for Teacher Expectations (Model 3)

Response	Teacher expectations	
	<i>B</i>	<i>SEb</i>
Fixed Part		
Intercept	4,150*	0,015
School leavers test	0,082*	0,002
Language test	0,005*	0,001
Mathematics test	0,015*	0,001
Reading comprehension test	0,008*	0,001
Gender (girl)	0,013	0,012
Ethnicity (minority)	0,055*	0,017
Parental education (Low)	-0,108*	0,016
Parental education (High)	0,085*	0,016
Ethnicity (minority) * School leavers test	0,002*	0,001
Parental education (Low) * School leavers test	-0,005*	0,001
Parental education (High * School leavers test	-0,002	0,001
Closeness	0,021	0,016
Closeness * Gender (girl)	0,006	0,018
Closeness * Ethnicity (minority)	0,034	0,021
Closeness * Parental education (Low)	-0,022	0,022
Closeness * Parental education (High)	0,001	0,021
Closeness * School leavers test	0,000	0,001
Conflict	-0,010	0,015
Conflict * Gender (girl)	0,013	0,018
Conflict * Ethnicity (minority)	-0,003	0,022
Conflict * Parental education (Low)	0,009	0,020
Conflict * Parental education (High)	0,013	0,021
Conflict * School leavers test	0,000	0,001
Dependency	-0,081*	0,017
Dependency * Gender (girl)	0,012	0,017
Dependency * Ethnicity (minority)	0,000	0,021
Dependency * Parental education (Low)	-0,003	0,023

Dependency * Parental education (High)	0,016	0,022
Dependency * School leavers test	-0,003*	0,001
Random Part		
<i>Class level</i>		
Intercept variance	0,048	0,004
Conflict slope variance	0,002	0,002
Dependency slope variance	0,003	0,002
Intercept-conflict covariance	0,005	0,002
Intercept-dependency covariance	0,002	0,002
Conflict-dependency covariance	0,002	0,002
<i>Student level</i>		
Intercept variance	0,236	0,004
Model fit		
-2*loglikelihood	14709.184	

Note. Number of classes = 614. Number of students = 9,881. * $p < .05$

Table 7

Indications of between class differences by means of 95% coverage intervals

	Variance	Standard error	Fixed coefficient	Lower limit 95% CI	Upper limit 95% CI
Intercept variance	0.048	0.219	4.150	3.721	4.579
Conflict slope variance	0.002	0.045	-0.010	-0.098	0.078
Dependency slope variance	0.003	0.055	-0.081	-0.189	0.027