Journal of Applied Developmental Psychology 35 (2014) 181–191

Contents lists available at ScienceDirect

Journal of Applied Developmental Psychology

Successive teacher expectation effects across the early school years

Christine M. Rubie-Davies [a,*], Rhona S. Weinstein [b], Francis L. Huang [c], Anne Gregory [d], Philip A. Cowan [b], Carolyn P. Cowan [b]

[a] Faculty of Education, The University of Auckland, Private Bag 92601, Auckland, New Zealand
[b] Department of Psychology, University of California at Berkeley, Tolman Hall #1650, Berkeley, CA 94720-1650, USA
[c] Educational, School, and Counseling Psychology, University of Missouri, 16 Hill Hall, MO 65211, USA
[d] Graduate School of Applied and Professional Psychology, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854-8085, USA

ARTICLE INFO

Article history:
Received 20 December 2012
Received in revised form 13 March 2014
Accepted 18 March 2014
Available online xxxx

Keywords:
Teacher expectations
Long-term effects
Teacher beliefs
Academic achievement

ABSTRACT

The capacity for teacher expectation effects to interact and compound across a child's schooling offers a largely untested mechanism for magnifying or minimizing effects. This study examined four types of long-term teacher expectation effects: within-year effects of single teachers, cross-year effects of single teachers, mediated effects of single and multiple teachers, and compounded effects of multiple teachers. Participants were 110 students tracked from preschool through Grade 4 on measures of achievement and teacher expectations. Evidence was found for within-year but not direct cross-year effects. However, path models demonstrated enduring indirect effects of teacher expectations on cross-year achievement. Multiple years of teacher expectation effects were additive in predicting student achievement at fourth grade, with similar effects for teachers' over- and underestimates of student ability. The study extends understanding of longer-term teacher expectation effects.

For over four decades, since the classic Rosenthal and Jacobson (1968) study, the voluminous research on teacher expectations has shown, in both experimental and correlational studies, that the self-fulfilling prophecy effect does exist in classrooms (see meta-analyses and reviews by Babad, 2009; Brophy, 1983, 1985; Good & Weinstein, 1986; Harris & Rosenthal, 1985; Hattie, 2009; Jussim & Harber, 2005; McKown, Gregory, & Weinstein, 2010; Raudenbush, 1984; Spitz, 1999; Weinstein, 2002). That is, teacher expectations, the beliefs that teachers hold about the potential academic performance of their students, *can* become confirmed in reality. However, in order for the expectations of teachers to have impact on students, they must be expressed in some way. Changes in student performance are hypothesized to result from differential interactions with teachers, which provide disparate learning opportunities for students for whom teachers hold high or low expectations and/or which communicate messages to students about differential ability. Both opportunities to learn and messages about ability can have an impact on student motivation and learning (Brophy & Good, 1974; Weinstein, 2002). There is growing evidence about such mediating processes between teacher expectation and student outcome, in the form of specific teacher behaviors that bring about such effects and in the form of student awareness of differential teacher treatment that signals ability differences. There is also evidence about moderating factors that magnify or lessen effects, such as differential susceptibility to producing such effects (in teachers) and to responding to such effects (in students).

Controversy continues, however, about the size of teacher expectation effects. Some researchers have argued that the effects of teacher expectations on student achievement outcomes within a single school year are, on average, small, resulting in a 5–10% difference in student achievement (Brophy, 1983), whereas other researchers, who have measured moderators such as the time of year of the expectation manipulation or teacher differences such as the level of differential treatment in the classroom or the positivity of class-level expectations, have reported much larger effects (see Bohlmann & Weinstein, 2013; Brattesani, Weinstein, & Marshall, 1984; McKown & Weinstein, 2008; Raudenbush, 1984; Rubie-Davies, 2007). Continued contextual analysis of expectation processes and their effects is crucial to advance understanding of the conditions under which this social influence phenomenon is operative (Weinstein, 2002).

Despite the large body of research on teacher expectation effects, most studies have been conducted within a relatively short time frame, one year or less. Relatively less is known about the longer-term effects either of a single teacher or of multiple teachers. In the sparse research literature available on longer-term relations between teacher expectations and student achievement, no studies have explored how expectation effects, occurring with different teachers over multiple school years, can interact and compound over time. Expectation effects could potentially become more powerful when viewed through a longitudinal lens over the course of a student's achievement history. The capacity for such effects to interact and compound across a child's school career offers one viable and largely untested mechanism for magnifying

* Corresponding author at: School of Learning, Development and Professional Practice, Faculty of Education, The University of Auckland, Private Bag 92601, Auckland 1150, New Zealand. Tel.: +64 9 373 7599x82974; fax: +64 9 623 8827.
E-mail address: c.rubie@auckland.ac.nz (C.M. Rubie-Davies).

http://dx.doi.org/10.1016/j.appdev.2014.03.006
0193-3973/

or minimizing such effects, critical to the debate about the strength of teacher expectation effects.

We review this literature in relation to various conceptualizations of the longer-term effects of teacher expectations on student outcomes. We examine the evidence for teacher expectation effects that carry over across school years, both at the elementary and secondary levels. We then address the evidence for the accumulation or dissipation of such teacher expectations effects, most commonly tested as the lingering effects of a single teacher across time. Finally, we suggest a reframing of the "accumulation" question to include the study of multiple teachers. Given the gaps in this literature, the current study investigated the longer-term effects of both single and importantly, multiple teacher expectations on student achievement across five years, from kindergarten to the end of fourth-grade.

Cross-year teacher expectation effects

Although most expectation studies examine within-year expectation effects, there has been growing interest in the *cross-year* effect (i.e., carryover, durability, and sustainability) of a given teacher on future outcomes in subsequent school years. Evidence for cross-year effects exists both at the elementary and secondary level. Rosenthal and Jacobson (1968) found an expectation advantage for the induced intellectual bloomers for the fifth graders that persisted into the second year. Further, although achievement was not measured, Rist (1970) documented a potential mediating factor of cross-year effects in the relatively fixed nature of reading group assignments in terms of curricular exposure and labeling that persisted from kindergarten to second grade.

A study by Alvidrez and Weinstein (1999) demonstrated that preschool teachers' over- and underestimates of children's intelligence at age 4, relative to measured IQ, predicted grade point average (overall beta weight of almost .4) and taking of the Scholastic Aptitude Test 14 years later. Overestimates are defined as teacher expectations which are higher than prior student-measured intelligence or achievement would predict while underestimates are expectations that are lower relative to student IQ or achievement. Thus, beyond the effects of early IQ, teacher expectations of preschoolers' intelligence could predict academic outcomes as students were entering college. Of import, these effects were moderated by type of expectation and quality of the home environment. Specifically, teacher predictions were strongest for underestimated children and weakest for children whose homes were more educationally-oriented. That is, teacher expectations had greater effects on children for whom expectations were low relative to achievement and lesser effects when students came from a home background of rich educational experiences. In a similar study by Sorhagen (2013) across 10 national sites, teachers' over- and underestimates of student achievement in first grade predicted student achievement at age 15. Students from low income backgrounds were most vulnerable to teacher expectations, particularly when their mathematics and language abilities were underestimated (with effects less in reading).

Hinnant, O'Brien, and Ghazarian (2009) found that early teacher expectations at first and third grade predicted child mathematics but not reading performance at fifth grade. In a European study, Gut, Reimann, and Grob (2013) showed that both parent and teacher expectations of children's competence at ages 5–7 predicted academic performance three years later. Of interest, the higher the family adversity and children's behavior problems, the lower the expectations of child competence by parents and teachers and expectations mediated the relation between risk factors and future child performance.

Similarly, at the secondary level, in a study of students from sixth to twelfth grades, Smith, Jussim, and Eccles (1999) documented that seventh grade, but not sixth grade, teacher expectations (in this study defined as perceptions of performance, talent, and effort) predicted the number of nonremedial mathematics courses that students took in high school. For every standard deviation increase in teacher perceptions, students enrolled in an average of 0.25 more mathematics courses in high school. In a sample of ethnically diverse youth aged 6 through 16, Mistry, White, Benner, and Huynh (2007) demonstrated a cross-year effect three years later of teachers' expectations on GPA. In the Dutch context, de Boer, Bosker, and Van der Werf (2010) showed a strong relation between early bias in teacher expectations (the difference between expectations and actual performance) at entry and later student achievement in the fifth year of secondary school. Student academic outcomes after five years were lowest for students whose teacher had a severe negative expectation bias, with a difference in achievement of approximately one full school year, a substantial effect.

These studies underscore that after controlling for the prior achievement or ability of students, *earlier* teacher expectations can have lasting cross-year effects (shorter-term and longer-term) on *later* outcomes, such as achievement, course-taking, and test-taking for college admission. Importantly, context (e.g., moderating effect of home environment in Alvidrez & Weinstein, 1999) and domain (e.g., mathematics but not reading in Hinnant et al., 2009) have been shown to be critical. This predictive ability of teacher perceptions beyond one school year and up to 14 years, at both elementary and secondary levels, is an important phenomenon in its own right.

Do teacher expectation effects accumulate or dissipate?

Given findings for enduring cross-year effects of a single teacher's expectations, researchers have pressed to quantify these lasting links as evidence for or against the strength of expectation effects. Smith et al. (1999) argued that such cross-year effects of single teachers could become stronger, remain stable, or dissipate across time. These authors introduced the term "accumulation" to the expectation literature and defined it as follows: "that a self-fulfilling prophecy triggered at one time exerts an increasingly larger influence over targets as time passes (p. 548)." In contrast, dissipation represents a decreasing effect of the original teacher's expectations on student outcomes over time. Examinations of patterns of beta weights (reflecting the relation between biased earlier teacher expectations and later student achievement) across time and within multiple contexts (single teacher within one school year, a single teacher across subsequent years, and multiple single teachers across multiple years) have yielded conflicting findings, with evidence for dissipation or weakening of effects as well as for stability of effects over time (de Boer et al., 2010; Hinnant et al., 2009; Jussim & Harber, 2005; Jussim, Robustelli, & Cain, 2009; Smith et al., 1999).

Accumulation reframed: multiple teachers, mediating processes, and compounded effects

A deeper understanding of accumulation of expectation effects must look at the dynamics beneath the cross-year influence of a single teacher. This includes a consideration of the interrelations between the expectation effects of multiple teachers (and the compounding of effects) for the same students across school years — both direct and indirect pathways that lie between the first teacher expectation and future student achievement. Brophy (1983) was the first to suggest that even if teacher expectation effects were small in one year, their effects could increase markedly as they compounded across years.

This compounding of effects across multiple teachers and across school years has largely been untested. One exception was a study by Blatchford, Burke, Farquhar, Plewis, and Tizard (1989) that examined the effect of two years of teacher expectations on progress across the three years of infant school in the UK. Progress was defined as "relative change over the year for children with equal scores at the start of the year" (p. 26). By the end of three years, the size of the effect for expectations (an overall rating across three years), even controlling for curriculum coverage, was 0.9 standard deviation units for both mathematics and language, a sizeable effect. Thus, in contrast to Smith et al. (1999), our conceptualization of accumulation revisits this observation of Brophy (1983). It moves beyond a single teacher toward the examination of *annual expectation effects* across a child's achievement history, from the first to subsequent teachers.

The increasing size of expectation effects across years rests upon a similar pattern of biased perception (as well as actions taken) by successive teachers across school years. That is, students are repeatedly exposed to the same pattern of over- or underestimates of ability (relative to prior performance) as well as to teachers prone to producing expectancy effects (Rubie-Davies, 2008; Weinstein, 2002). It is an empirical question whether every teacher would see a student in exactly the same manner across time. Indeed, Hinnant et al. (2009) found modest correlations on average between the expectations of first-, third-, and fifth-grade teachers, ranging from .24 to .10. Studies have also identified teacher differences that moderate the strength of expectation effects, such as the propensity for biased expectations (Babad, 2009), high expectations for all students (Rubie-Davies, Hattie, Townsend, & Hamilton, 2007), and differential treatment towards those expected to achieve at high and low levels (Brattesani et al., 1984; McKown & Weinstein, 2008). Hence, students can be subjected to differing expectations as well as expectation effects across their school career that may alter the course of their achievement trajectory. A stronger test of accumulation of effects would utilize the same sample of students followed longitudinally and examine interrelations between expectation variables and the compounding of expectation effects produced each year for each student across school years.

Also especially important to assess are the earliest measures of student ability prior to the onset of school and the potential influences of teacher expectations on students' performance (Alvidrez & Weinstein, 1999). The first teacher can be regarded as setting in place expectation effects that may endure for several years of a student's schooling. Expectations when students enter school are based on a variety of factors such as social class (Rist, 1970), ethnicity (McKown & Weinstein, 2008; Rubie-Davies, Hattie, & Hamilton, 2006), and gender (Palardy, 1969; Tizard, Blatchford, Burke, Farquhar, & Plewis, 1988), as well as other attributes gleaned from early interactions with the child. Teachers have little achievement evidence on which to base their expectations of students on school entry and hence biased expectations at this stage may have important and larger effects on student outcomes than we might expect to find in later years (Kuklinski & Weinstein, 2001).

The present study

Hence, in this longitudinal study beginning preKindergarten, we expand on current research on longer-term expectation effects by examining interrelations between the expectations of single and multiple teachers and achievement for the same students across time. Our aims for this study were: a) to examine the effects of a single teacher's expectations on student achievement within one school year (*within-year teacher expectation effects*); b) to investigate the enduring direct effects of a single teacher's expectations on student achievement across school years (*cross-year teacher expectation effects*); c) to identify the enduring indirect pathways through which single teachers' expectations affect subsequent student achievement, and subsequent teachers' expectations over time (*mediated teacher expectation effects*); and d) to explore the *compounding* of annual teacher expectation effects and the relative effects of over- versus underestimation of ability on fourth-grade achievement for the same child with multiple teachers across school years (*compounded teacher expectation effects*).

Method

This study used data from the Schoolchildren and their Families Project (Cowan, Cowan, Ablow, Johnson, & Measelle, 2005), which began in 1990 and was based in the southwestern area of the United States. The project recruited families from 28 cities and towns to participate in a study focused on children's transition to kindergarten, particularly on family relationship factors that enhanced or interfered with the children's ability to make a successful transition. Also assessed were teachers' perceptions of students in a variety of domains, such as academic capability, from which a measure of teacher expectations was constructed (further outlined below).

In all, 110 families joined the study. Data were collected before the children entered school, and then at Kindergarten and Grades 1, 4, 9, and 11. Given some subject attrition after Grade 4, this study focused on the period between preKindergarten and end of Grade 4 for which there was no attrition and consistent measurement of teacher perceptions. The students included in the current study were self-selected through their parents opting into the Schoolchildren and their Families Project. Approximately two-thirds of the parents in the sample were involved in a group intervention[1] with a parenting or a marital emphasis, prior to the transition into kindergarten.

Participants

The study incorporated measures related to two groups of participants: the sample of 110 students (the families' first children; 64 boys and 46 girls), who were followed from preschool through the end of Grade 4, and the teachers of these students in Kindergarten (K), Grade 1 (G1), and Grade 4 (G4). More affluent families were overrepresented in the sample with a median total family income of $80,000 in 1990. However, the standard deviation of approximately $62,000 suggests a wide dispersion of income among the sample. With regard to ethnicity, 89 students were European American, 10 were Asian or Middle Eastern American, 6 were African American and 2 were Hispanic (for 3 students no ethnicity data were available). At K, G1, and G4, teachers provided expectation data for the target student in their class. The teacher sample ($N = 287$) consisted of 103 K teachers, 94 G1 teachers and 90 G4 teachers. Of the original 110 students, only 4 were ever enrolled in the same class as another target child in the study at any time point. Hence, of the 110 students in the study × 3 data points (K, G1 and G4) (330 data points), only 4 pairs represent non-independent measures. This indicates that there was extremely minimal "nesting" or cross classification of multiple students across shared classrooms over the five years.

Measures

Student characteristics

A number of student characteristics were considered as possible covariates in the statistical models including student socioeconomic status, race/ethnicity, gender, and preschool attendance. Parents reported on their child's gender, race/ethnicity, and preschool enrollment. They also shared their own level of education. There is evidence that mother's education is a more reliable indicator of socioeconomic status than total family income (Marks, 2008) and hence, mother's level of education was a proxy for socioeconomic status in this study. Level of education was categorized on a 7-point scale with 1 indicating the lowest level

[1] The parents of these students were participants in a preventive intervention study that examined the effects of a parenting and marital intervention prior to the first child's entry into kindergarten. The intervention focused on improving the quality of the couples' interactions as partners and as parents of their children (Cowan, Cowan, Ablow et al., 2005). The parents were not aware of the possibilities of intervention when they entered the study and the intervention was offered randomly once parents expressed interest in joining the study (i.e., this was not a sample recruited on the basis of family relationship distress). About two-thirds of students were assigned a "1" as their parents were in the intervention condition and one-third of students were assigned a "0" as their parents were in the comparison condition. Prior research showing the contrast between intervention and control participants revealed minimal direct effects of the intervention on academic achievement (Cowan, Cowan, & Heming, 2005). We included intervention condition as a covariate in preliminary analyses. We modeled results using a binary indicator of intervention status (1 = yes/0 = no) and results related to intervention status were all not statistically significant.

of schooling in this sample (some high school), 4 indicating a college graduate, and 7 indicating the highest (doctorate or MD). On average, the mothers attained high levels of education ($M = 5.34$, $SD = 1.1$), with the mean education beyond the completion of a bachelor's degree. Overall, 11% had a doctorate or MD, 37% had a master's degree, 40% were college graduates, and 12% had less than a college education.

Student verbal ability and achievement

Standardized measures of verbal ability and achievement were obtained from students at age 4 and at the end of each school year. Specifically, at age 4, students completed the Peabody Picture Vocabulary Test (PPVT), which provides an estimate of the child's preK verbal ability. Raw scores are converted into standard scores based on age, with a mean of 100 and a standard deviation of 15. Students also completed the Peabody Individual Achievement Test (PIAT) at the end of K, G1, and G4 (Markwardt, 1989). Researchers administered the tests to each student during home visits, which ensured that teachers did not have access to the scores. The PIAT assesses achievement in general knowledge, reading recognition, reading comprehension, mathematics knowledge and application of concepts and facts, and recognition of correct spelling. A full-scale score was derived from the mathematics and reading subscales and used as an index of academic achievement.

Teacher expectations of individual students

Similar to previous work (e.g., Alvidrez & Weinstein, 1999; Donohue, Weinstein, Cowan, & Cowan, 2000), we operationalized teacher expectations as a teacher estimate of the child's academic capability. The subscale of the *Child Adaptive Behavior Inventory* (*CABI*) entitled "intelligent" (Cowan, Cowan, Heming, & Miller, 1995) was used as a measure of teachers' expectations because the items reflected teachers' *perceptions* of student ability. This subscale consists of 5 items, scored from 1 = *Not at all like this child* to 4 = *Very much like this child*. Sample items included: "This child is smart for his/her age," "This child understands difficult words." Reported Cronbach alpha reliability for this subscale was $\alpha = .81$ for K, $\alpha = .85$ for G1, and $\alpha = .89$ for G4 (Cowan, Cowan, Ablow et al., 2005; Cowan, Cowan, & Heming, 2005). High scores on the intelligence scale indicated that the teacher considered the student academically capable. It is of note that the expectations for the children in this sample were generally high (approximately 3.4 at each grade as reported in Table 1, with a minimum score of 2.27 and a maximum of 3.98) which is possibly not surprising given the mean for mothers' education.

Procedures

In the Fall of K, G1, and G4, teachers completed the teachers' expectations (Cowan, Cowan, Ablow et al., 2005; Cowan, Cowan, & Heming, 2005) for the target students in their respective classes. Spring ability/achievement scores for students were available for preK, K, G1, and G4. This enabled a test of the temporal relations between Fall teacher ratings (collected in October each year and representing early expectations of students) and Spring achievement with prior year ability/achievement as a control — that is, a test of teacher expectation effects.

Table 1
Means and standard deviations for verbal ability/achievement, teacher expectations and proportions for student characteristics.

	%	n	M	SD
PreK verbal ability		78	120.21	14.01
K achievement		104	106.68	15.90
K expectations		97	3.47	0.50
G1 achievement		95	114.17	17.80
G1 expectations		95	3.40	0.57
G4 achievement		93	118.54	14.09
G4 expectations		98	3.38	0.52
Female students	40%			
Attended preschools	95%			

Results

Predicting within-year, cross-year, mediated, and compounded expectation effects

Data analytic approach

We examined four research aims concerning: 1) Within-year expectation effects, 2) Cross-year expectation effects, 3) Mediated expectation effects, and 4) Compounded teacher expectation effects using a cross-lagged panel design (CLPD) to test the compounding of teacher expectation effects. The within-year expectation effects are defined as the effects of a single teacher within a particular grade; cross-year effects are single teacher expectation effects of an earlier teacher on student achievement several years later; mediated teacher expectation effects trace the influence of single teachers on later achievement through intervening variables such as the expectations of subsequent teachers over time; and compounded teacher expectation effects reflect both direct and indirect pathways of multiple teachers in predicting future achievement. CLPD takes advantage of the temporal order of events and can be used to study the reciprocal relation of several variables (Campbell & Stanley, 1963; Kenny, 1975, 2005). For example, prior teacher expectations may influence student achievement at a future time, which then can influence teacher expectation again in a succeeding time period. CLPD has been used to specifically study the causal influence of teacher expectations on academic performance (e.g., Crano & Mellon, 1978) or how student attitudes affect performance, which then affects future attitudes (Ma & Xu, 2004). Given the correlational nature of the data, theorized causal links can be examined, yet definite causal claims cannot be made (e.g., Rogosa, 1980).

CLPD makes use of multiple regression and path analytic techniques, which are used to estimate the standardized path coefficients between variables and are used to model both direct (denoted by an arrow from one variable linked directly to another variable) and indirect associations (association is through another mediating or intervening variable) between variables. Instead of running several separate regressions, a single CLPD can model the relation among variables in a straightforward and parsimonious manner.

Missing data

Prior to conducting our analyses, we inspected the data for missing data. Although there was no attrition across the sample, there were data missing for some students at some time points. The largest proportion of the sample (29%) was missing scores on the PPVT. We conducted Little's (1988) test to assess whether data were missing completely at random (MCAR), which would allow simple listwise deletion of data and not bias results. Results indicated however that data were not MCAR ($\chi^2[206] = 254.72, p < .05$). While there is no established cutoff criterion for what is deemed an acceptable percentage of missing data to yield valid estimates (Dong & Peng, 2013), we followed practical guidelines set by Allison (2012) and Bodner (2008). To account for missing data, we used 30 multiply-imputed datasets in all the analyses. Multiple imputation (Rubin, 1987; Schafer & Graham, 2002) is robust to various types of missing data mechanisms (McKnight, McKnight, Sidani, & Figueredo, 2007; Schafer, 1999).[2] The imputed datasets were created using SAS PROC MI using all variables in the dataset as auxiliary variables (e.g., race/ethnicity of parents).

[2] As a robustness check, we also ran the CLPD models using Full Information Maximum Likelihood (FIML), an alternative procedure for handling missing data. Results using multiply-imputed datasets and FIML were very similar (cf. Enders, Dietz, Montague, & Dixon, 2006); all statistically significant and nonsignificant paths remained the same with the exception of the association of gender with grade 4 achievement (using FIML, $B = -.14$, $p = .03$).

In all the analyses, we used robust standard errors (MLR) which were robust to nonnormality as some of the teacher expectations variables were slightly skewed (*skew* = −1.07 to 0.87). However, in addition to accounting for slight nonnormality, skewness values were well within acceptable limits for our analyses (Lomax & Hahs-Vaughn, 2012). Analyses of imputed datasets were done using Mplus (Muthén & Muthén, 2010), which has built in procedures to model complex (i.e., multiple mediated paths) indirect effects and compute standard errors using the Delta method (Sobel, 1987), allowing for tests of statistical significance for mediated pathways.

After examining the full model, we trimmed nonsignificant path coefficients from the teacher expectation variables to arrive at a final, more parsimonious model. We present standardized path coefficients, which can be interpreted as standardized regression coefficients. The total association (or 'effect' as it is commonly referred to in path models) of one variable with another variable is the sum of the direct and indirect effects.

To assess model fit, we used various fit indices to evaluate the quality of the resulting model (Fan & Sivo, 2007). Aside from using χ^2 as a measure of model fit where nonsignificant values indicate good fitting models, we used additional measures of model quality such as the Tucker–Lewis Index (TLI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). CFI and TLI values greater than .95 are considered good fitting models (Hu & Bentler, 1999) while for RMSEA, values less than .08 are considered reasonable (Kline, 2005).

Descriptive analyses

We examined the data descriptively with regard to the consistency of student achievement over time and the consistency of teacher expectations over time. Student academic test results in successive years were all positively and significantly correlated (see Table 2). Correlations between the preK verbal ability test and the achievement test results were moderate at K ($r = .41$), G1 ($r = .44$), and G4 ($r = .56$). Despite an apparent increase in the size of the correlation from K to G4, the differences between these correlations, using a Fisher r to z transformation for comparison, were not significant. Larger correlations were seen between the achievement tests at various grade levels (between K and G1, $r = .78$; between K and G4, $r = .60$; and between G1 and G4, $r = .75$). The K–G4 correlation was significantly different from the K–G1 ($p = .04$) and the G1–G4 correlations ($p = .02$). These correlations reflect the relative stability of achievement measures, yet even with the highest correlation between K and G1 (.78) only 61% of the variance was explained by prior achievement, leaving room for other contributing factors.

Pearson correlations were calculated to investigate the consistency of teacher ratings for student academic capability in the fall of each year, from K to G4. Teacher expectations at the three grade levels were positively and significantly correlated: between K and G1, $r = .46$, $p < .001$; K and G4, $r = .25$, $p < .01$; G1 and G4, $r = .26$, $p < .01$. As can be seen, the correlation between K and G1 teachers' expectations was moderate across the span of one year, while the correlations between K and G4, and G1 and G4 teachers (for time periods of four and three years, respectively) were small, showing that there was variation in how the same students were viewed by different teachers.

Parsimonious model

We ran an initial set of CLPD models (not shown) with all the covariates to establish which set of covariates needed to be maintained in the final model. Gender and preschool attendance were maintained in the model given their significant relation with student achievement (e.g., Chiu & McBride-Chang, 2006; Huang, Invernizzi, & Drake, 2012). Whether or not the student had parents who participated in the intervention had no significant direct pathways in the model. Student SES and race also were not significantly related to teacher expectations or PreK and G1 achievement. Moreover, when intervention condition, SES, and race were included in the model, the pattern of findings among expectations and achievement did not change. Thus, for the sake of parsimony, student SES, race, and intervention condition were not included as covariates.

The initial hypothesized model fit the data well based on all model fit indices, χ^2 (11) = 11.98, $p = .37$, *RMSEA* = .04, *CFI* = .99, *TLI* = .98. However, three path coefficients from the various teacher expectation variables were not statistically significant (*ps* > .05), which indicated that the model could be trimmed slightly. More specifically, all of the associations of the prior year teacher expectations (cross-year teacher expectation effects) did not have statistically significant (all *ps* > .05) associations with future achievement (see Fig. 1). Specifically, K expectations did not have a direct association with G1 achievement ($\beta = 0.10$, $p > .05$) and G4 achievement ($\beta = -0.04$, $p > .05$). Similarly, G1 expectations did not have a direct association with G4 achievement ($\beta = -0.01$, $p > .05$).

A revised CLPD model was run, which is often done in exploratory path analytic models (Schreiber, Nora, Stage, Barlow, & King, 2006; Stage, Carter, & Nora, 2004); the revised model removed the three nonsignificant path associations of prior year teacher expectations with future achievement (see Fig. 2). The resulting model had excellent fit indices as well, χ^2 (14) = 17.09, $p = .25$, *RMSEA* = .04, *CFI* = .99, *TLI* = .98. Removing nonsignificant variables that are correlated with other independent variables in the model has the effect of improving model power (i.e., reduces collinearity among variables and improves the precision of standard errors) and results in a more parsimonious model. One difference with the initial and revised model is that G1 teacher expectations were now statistically significantly related to G1 achievement in the revised model ($\beta = 0.15$, $p = .02$).

Model R^2 for each endogenous variable (i.e., variable receiving effects or dependent variables) are presented in Table 3. R^2 offers a measure of the variance explained for a given variable based on significant direct paths. For example, G1 teacher expectations R^2 (.29***, $p < .001$) is the variance explained when G1 teacher expectations were regressed on two variables (K expectations and K achievement). In other words, 29% of the variance in G1 teacher expectations was explained by prior kindergarten achievement and prior teacher

Table 2
Correlations among student achievement/ability, teacher expectations, and student characteristics.

	2.	3.	4.	5.	6.	7.	8.	9.
1. PreK verbal ability	.41***	.44***	.56***	.20*	.18	.09	−.12	.00
2. K achievement	–	.78***	.60***	.28**	.40***	.26**	−.20*	.08
3. G1 achievement		–	.75***	.37***	.42***	.38**	−.16	.19*
4. G4 achievement			–	.27**	.29**	.45***	−.25**	.21*
5. K teacher expectations				–	.46***	.25**	.03	.13
6. G1 teacher expectations					–	.26**	.07	−.03
7. G4 teacher expectations						–	−.02	.19*
8. Gender, female (1) male (0)							–	−.02
9. Preschool attendance								–

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

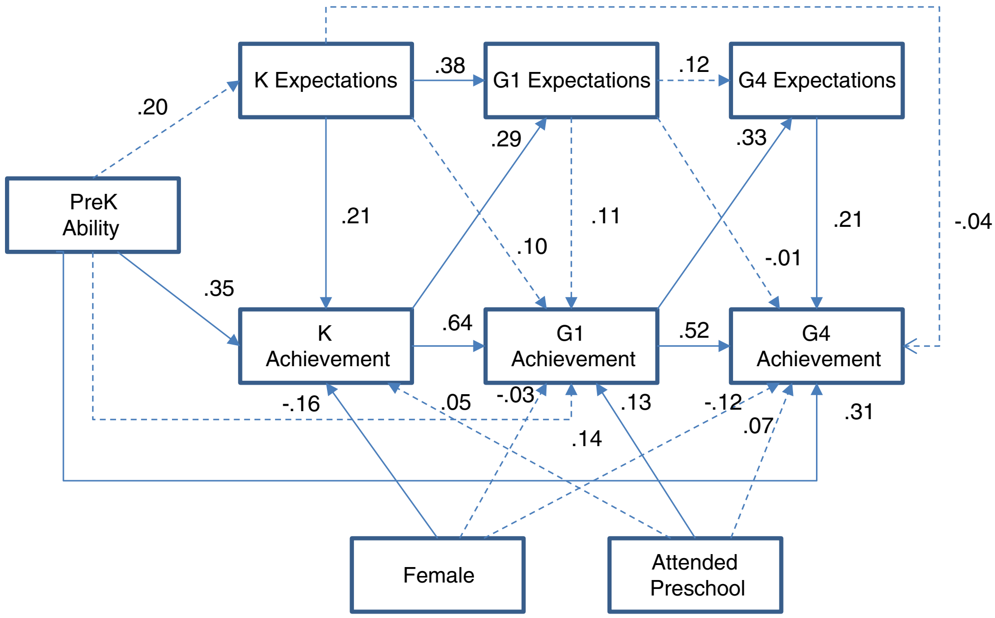

Fig. 1. Standardized path coefficients shown for hypothesized model ($n = 110$). Solid lines indicate statistically significant associations ($p < .05$). Dotted lines indicate nonsignificant results ($p > .05$).

expectations. In contrast, it is noteworthy, that no significant variance in K teacher expectations was explained by previous achievement (PreK verbal ability). Further, over a longer span of three years, G1 achievement, but not G1 teacher expectations, explained 16% of the variance in G4 teacher expectations (see Table 3).

Within-year expectation effects

In the trimmed model, all of the within-year teacher expectations had a positive and statistically significant association with student achievement (K expectation to K achievement $\beta = 0.21$, G1 expectation to G1 achievement $\beta = 0.15$, G4 expectation to G4 achievement $\beta = 0.21$, all $ps < .05$). In other words, the higher the teacher expectations, the greater the student achievement, after accounting for prior abilities (i.e., preK verbal ability and/or prior year PIAT), gender, and preschool attendance. One standard deviation increase in K, G1, or G4 expectations was associated with 0.15 to 0.21 standard deviation increase in student achievement scores. Importantly, as tested more stringently within a temporal path model, teacher expectation effects at Grades 1 and 4 were significant *over and above* the expectation effects of the prior year(s). This suggests an additive pattern of successive expectation effects in G1 and G4, beyond the effects documented in kindergarten.

Cross-year expectation effects

The initial CLPD model showed that teacher expectations in a single year did not predict achievement in future years, suggesting no *direct* cross-year expectation effects (Fig. 1). As mentioned previously, accounting for student characteristics and early verbal ability, the

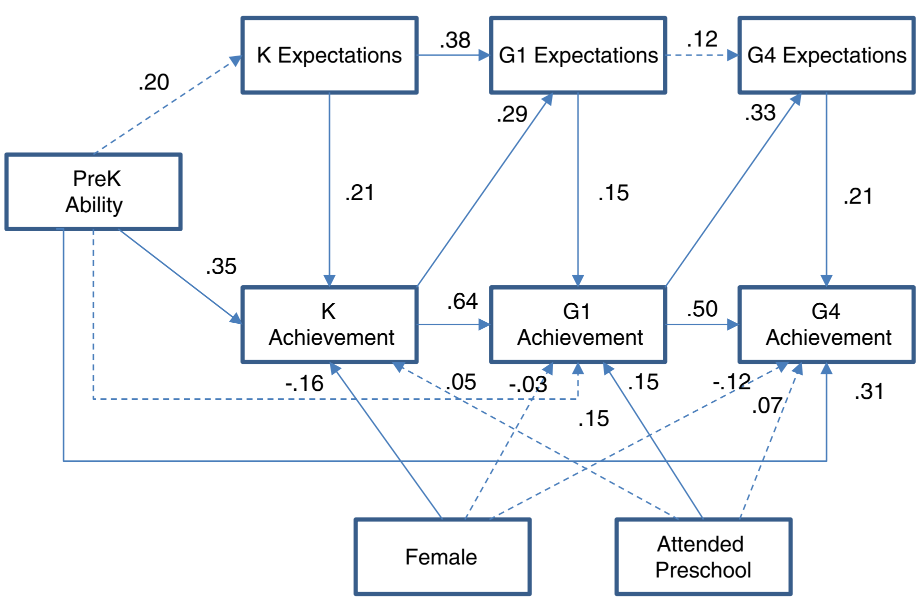

Fig. 2. Standardized path coefficients shown for revised model ($n = 110$). Solid lines indicate statistically significant associations ($p < .05$). Dotted lines indicate nonsignificant results ($p > .05$).

Table 3
Proportion of variance accounted for per endogenous variable ($n = 110$).

Endogenous variable	R^2
K expectations	.04
G1 expectations	.29***
G4 expectations	.16*
K achievement	.24**
G1 achievement	.66***
G4 achievement	.70***

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

Table 4
Summary of types of expectancy effects ($n = 110$).

	Model pathways	Standardized estimate
Within-year	Kexp ➔ Kach	.21** (.08)
	G1exp ➔ G1ach	.16* (.07)
	G4exp ➔ G4ach	.23*** (.06)
Cross-year (direct effects)	Kexp ➔ G1ach	*ns*
	Kexp ➔ G4ach	*ns*
	G1exp ➔ G4ach	*ns*
Mediated effects (indirect effects)	Kexp ➔ mediators ➔ G1ach	.20** (.06)
	Kexp ➔ mediators ➔ G4ach	.12** (.05)
	G1exp ➔ mediators ➔ G4ach	.11* (.05)

Note: Kexp = K Expectation, G1exp = G1 Expectation, G4exp = G4 Expectation; Kach = K Achievement, G1ach = G1 Achievement, G4ach = G4 Achievement; *ns* = not statistically significant. Standard errors are in parentheses.
* $p < .05$. ** $p < .01$. *** $p < .001$.

pathways from K teachers' expectations to G1 achievement, K teachers' expectations to G4 achievement, and G1 teachers' expectations to G4 achievement were non-significant (Fig. 1). In other words, the direct paths from single teacher effects to future achievement did not endure *across* school years. PreK verbal ability, female gender, G1 achievement, and G4 teacher expectations explained 70% of G4 achievement (see Table 3).

Mediated and compounded cross-year expectation effects

However, the trimmed CLPD model (see Fig. 2) also shows that single teacher expectations have *indirect* effects on future achievement beyond their effects within a single school year. Whereas interpreting the direct association of one variable to another is straightforward (e.g., a one standard deviation increase in K expectations is associated with a 0.21 standard deviation increase in K achievement), modeling indirect associations requires additional computations, especially when computing standard errors (in Mplus using the MODEL INDIRECT or CONSTRAINT command). In addition, multiple paths need to be considered.

For example, to calculate the total mediated effect of K teacher expectations on G1 achievement we consider the following: The first indirect pathway of K teacher expectations to G1 achievement is via K achievement. Specifically, K teacher expectations are directly associated with K achievement (.21), which is, in turn, related to G1 achievement (.64), resulting in an indirect effect of .13 (i.e., .21 × .64). The next indirect route is via G1 teacher expectations. Specifically, K teacher expectations are directly associated with G1 teacher expectations (.38), which are in turn related to G1 achievement (.15), resulting in a product of .06. Finally, K teacher expectations also have indirect effects via K achievement (.21), which links to G1 expectations (.29) and then leads to G1 achievement (.15). The product of the three estimates is .01. Summing the three products (.13, .06, and .01) shows that the indirect effects of K teacher expectations on G1 achievement are .20 ($SE = .06$, $p < .01$). Given that there are no direct effects between the two variables, K teacher expectations' total indirect effects on G1 achievement remain .20. One standard deviation increase in K teacher expectations would be associated with increased G1 achievement in the magnitude of .20 standard deviations.

Further, with calculations reflecting the complex pathways, K teacher expectations had a total indirect and importantly, an enduring association with G4 achievement ($\beta = 0.12$, $SE = .05$, $p < .01$). Again, using calculations reflecting the complex associations over time, G1 teacher expectations also had a statistically significant indirect association with G4 achievement ($\beta = 0.11$, $SE = .05$, $p = .02$). One standard deviation increase in K or G1 teacher expectations would be associated with increased G4 achievement in the magnitude of 0.11 to 0.12 standard deviations.

Table 4 presents a summary of the types of expectation effects estimated in the CLPD model (Fig. 2). None of the standardized regression coefficients show differences with each other that are statistically significant given the large standard errors. However, all of the beta coefficients are significantly different than zero.

Thus, the results show significant temporal evidence for within-year teacher expectation effects at K, G1, and G4 but not for the cross-year direct effects of single teachers, which some might argue reflects dissipation of effects of the K and G1 teachers on G4 achievement. However, the significance of the indirect effects of teacher expectations at K and G1 (beyond the variance predicted by student differences in preK verbal ability) points to a mediated influence of earlier teachers on G4 achievement, which endures. That is, the prior years' expectations show compounding only through indirect effects since there are no direct effects of teacher expectations that span more than one year. Further, both direct and indirect pathways have comparable effect sizes, work together to predict student achievement at G4, and capture the interaction and compounding of expectation effects across school years.

Predicting compounded expectation effects and the relative effects of compounded over- versus underestimated ability

While the CLPD results provided evidence for the compounding of successive years of teacher expectation effects—inclusive of indirect as well as direct pathways—regression analysis provided another way to test compounding. It used *independent* estimates of yearly teacher expectation effects for individual students and also enabled a test of the comparative effects of positively- versus negatively-biased teacher expectations. Thus, to address research aim four concerning the contribution of the *compounding* of annual teacher expectation effects (for the same child with multiple teachers) in predicting fourth-grade achievement, we used multiple regression analyses with the 30 multiply-imputed datasets to arrive at unbiased parameter estimates. We quantified how much a teacher's expectations over- or underestimated ability for a student in a given year. We regressed fall teacher expectations on the student's prior achievement or ability scores (i.e., predicting fall K teacher expectation using preK verbal ability; predicting G1 teacher expectations using K achievement; predicting G4 teacher expectations using grade one achievement). Using residual analysis, we computed over- or underestimation of ability as the difference between the obtained teacher expectation score and the predicted teacher expectation score (i.e., $Y - Y$). Teacher bias resulted if a student had positive (overestimation) or negative (underestimation) residuals. Residual scores close to zero indicated that the predicted and observed scores were almost the same (i.e., no bias). Residuals were then standardized ($M = 0$, $SD = 1$) for each year (i.e., K, G1, G4) and summed together to create an overall over-/underestimation (compounded) score for each student ($M = 0.00$, $SD = 2.07$, min $= -6.08$, max $= 4.09$). We then modeled students' G4 achievement scores on the compounded expectation scores while controlling for gender, preschool attendance, and preschool verbal ability. Finally, we tested the differential teacher effects of over- and underestimation scores by examining possible nonlinear effects between the scores and G4 achievement. We entered a second-degree polynomial (i.e., a squared term) of the over- and underestimation scores as a predictor of G4 achievement. This enabled us to understand if compounded effects of multiple years of underestimated ability had a larger magnitude of effect compared to multiple years of overestimated ability.

The regression analysis, using over- and underestimates of children's ability as perceived by teachers, tested the *compounding* (addition) of teacher expectation effects *for the same child* across school years with *multiple teachers* and compared the relative effects of over- versus underestimation (see Table 5). When entered as covariates in the first block, preK verbal ability ($\beta = 0.54$, $p < .001$) and preschool attendance ($\beta = 0.74$, $p < .001$) were both significantly associated with higher G4 achievement (Model A). The over- and underestimated scores were added in the next block and explained 5% additional variance in G4 achievement (R^2 change = .05, Model B). That is, the more teacher expectations over-estimated ability across the years (compounded), the greater the gains in achievement by 4th grade. Conversely, if student ability had greater underestimations over the years, students were more likely to have smaller achievement gains. The inclusion of the total over- and underestimation scores along with preK verbal ability, gender, and preschool attendance explained a total of 44% of the variance in G4 achievement.

The possible relative effects of over- versus underestimation on G4 achievement were tested in Model C. Results indicated that the squared term was not significant ($p = .15$). This shows that underestimates of student ability by teachers had a similar magnitude of effect on G4 achievement as overestimates of ability.

Discussion

The results of this study deepen our understanding about the accumulation of teacher expectation effects for individual students, as seen in the larger context of exposure to multiple teachers across the school careers of students. Importantly, as a window into the earliest days of formal schooling, this study begins with the first teacher in Kindergarten and with an assessment of verbal ability before students entered school at Kindergarten. We conceptualized four different long-term teacher expectation effects: within-year (single teacher), direct cross-year (single teacher), indirect or mediated cross-year (multiple teachers), and compounded cross-year (multiple teachers). With a sample of 110 students followed for five years from preKindergarten through fourth grade, we provided significant evidence for three of the four types of long-term effects of teacher expectations on student achievement.

First, looking within one school year, beyond the contribution of students' prior-year achievement, teacher expectations were found to independently and significantly predict students' year-end achievement at all three grade levels, at Kindergarten, first grade, and fourth grade. This predictive relation between early teacher expectations and later student achievement, beyond that explained by entering achievement differences, was seen *as early as* the Kindergarten year. Further, the significant teacher expectation effects at first and fourth grade each added unique predictive value beyond the expectation effects of previous years. These findings confirm the results of other studies (e.g., Blatchford et al., 1989; Kuklinski & Weinstein, 2001; Rubie-Davies, 2007) but also extend understanding by documenting the unique and additive expectation effects of subsequent teachers across school years.

Second, unlike the findings of previous work (e.g., Alvidrez & Weinstein, 1999; de Boer et al., 2010; Hinnant et al., 2009; Smith et al., 1999), we did not find *direct* predictive effects of single teachers across a number of school years. However, these teacher expectation effects did not disappear or dissipate, but rather, they were hidden in the dynamic interplay of indirect or mediated pathways. That is, factors at Kindergarten interacted with factors at first grade in ways that accumulated by fourth grade. The expectations of Kindergarten teachers did show predictive effects on first and fourth grade achievement that endured through mediated pathways. Similar indirect pathways were found for the influence of the expectations of first grade teachers on fourth grade student achievement. For every standard deviation increase in teachers' expectations, students' achievement scores increased by one-tenth to almost one-quarter of a point. This pattern of increase in achievement was comparable whether the teacher expectation effects reflected direct or indirect pathways.

Finally, the evidence for compounding of teacher expectation effects could be seen in two ways: in the significant pathways linking prior year's expectations with future achievement and in the addition of expectancy effects over multiple years. The model dynamically linked three successive years of teacher expectation effects to the prediction of student achievement in fourth grade. The regression results also demonstrated that beyond the effects of pre-Kindergarten ability, gender, and preschool attendance, the addition of three years of teacher expectation effects (for the same student and from multiple teachers) significantly predicted some of the variance in student achievement by fourth grade. The more that teachers overestimated the ability of students relative to prior achievement, the larger were the achievement gains by fourth grade — and this was among a group of children where achievement was a little above average, and where a ceiling effect could have been anticipated. Conversely, the more students' ability was underestimated across the grade levels, the lower their fourth grade achievement. Our finding of comparable effects for both over- and underestimation of student ability, however, did not replicate previous evidence for the greater effect of underestimated ability, such as in studies by Alvidrez and Weinstein (1999) who used the same methodology as this study and de Boer et al. (2010), who used a dummy variable to represent different degrees of over- versus underestimation of ability.

How do we make sense of these findings? As found in earlier work by Alexander and Entwisle (1988), the influence of the Kindergarten

Table 5
Predictors of fourth grade achievement ($n = 110$).

Variable	Model A *b*	Model A *SE*	Model A β^a	Model B *b*	Model B *SE*	Model B β^a	Model C *b*	Model C *SE*	Model C β^a
Intercept	46.66***	11.59		47.58***	11.22		50.64***	11.27	
Preschool verbal ability	0.54***	0.09	0.54	0.54***	0.09	0.54	0.53***	0.09	0.53
Female	−5.04*	2.33	−0.36	−5.81*	2.28	−0.41	−5.84*	2.28	−0.41
Attended preschool	10.39*	4.19	0.74	9.34*	4.08	0.66	8.60*	4.14	0.61
Over/under estimation				1.43*	0.55	0.21	1.07	0.58	0.16
Over/under estimation[b]							−0.27	0.20	−0.12
R^2	.39			.44			.45		
ΔR^2				.05			.01		

* $p < .05$. *** $p < .001$.
[a] Standardized regression coefficients for continuous variables. For dichotomous variables (i.e., Female and Attended preschool).
[b] Reflects a one standard deviation change when the variable = 1.

teacher may set students on a particular achievement trajectory but a trajectory that may be subject to some change along the way. In this sample, over and above the effect of Kindergarten teachers' expectations on student achievement, the first grade teachers nudged that trajectory, creating direct effects on first grade achievement and indirect effects on fourth grade achievement. Then the expectations of fourth grade teachers, through direct effects, also affected fourth grade student achievement. Thus, within-year expectation effects have longer-term consequences in part because achievement is correlated over time, prior achievement predicts later teacher expectations, and the expectations of multiple teachers are associated. A little “nudge” by the teacher in Kindergarten affected Kindergarten achievement, which in turn affected first grade achievement as well as importantly, the expectations of the first grade teacher. Of further interest, the expectations of these first teachers in the current study were not predicted by measures of children's preschool verbal ability, whereas differences in preschool ability predicted achievement in Kindergarten and continued to do so in fourth grade.

Evidence for some trajectory change can be found in the discontinuity of the relation between annual teacher expectation effects. Although these results show compounded expectation effects across years as hypothesized by Brophy (1983), they also point to a break in the links between the expectations of multiple teachers: that is, a significant link between the expectations of Kindergarten and first grade teachers but not between the expectations of first and fourth grade teachers. Variability in expectation correlations has been found by Hinnant et al. (2009), ranging between .24 and .10 across two years. This discontinuity in teacher expectations for students by fourth grade in this sample may reflect the longer time span (three years versus one year) and/or different criteria for perceived competence in the fourth grade transition. Hence, there appear to be two processes working in tandem; one is teacher effects (the same-year direct effects and the across-year indirect effects of different teachers' expectations, and presumably, treatment of students) and the other is student effects, whatever factors maintain consistency over time within students. Most importantly for the strength of expectation influence, consistency of teacher or student effects over time is critical.

Time and teacher expectation effects: future research

Future research on the compounding of expectation processes (assessing its strength, whether increasing, remaining stable, or dissipating) requires attention to effects that lie beneath the average of all students and all teachers. That is, we need to investigate subgroups of students whose abilities are consistently over- or underestimated across school years and subgroups of teachers who consistently produce expectation effects. The current sample was mostly white and middle class with mothers who were well educated. Hence the findings relate to a sample of students for whom expectation effects do not tend to be as strong (see Jussim, Eccles, & Madon, 1996). Therefore, future research among a broader socioeconomic group is warranted. Further, increasing size of effects rests on students being “subjected to the same or similar erroneous expectations over and over again” (Jussim & Harber, 2005, p. 146) by teachers as well as being subjected to similar differential treatment and it may be that the effects found in the current study would be greater with a more diverse population. Based on sociodemographic factors or course tracking, some students are at high risk for biased expectations of their ability and consistent differential treatment over time. For example, McKown and Weinstein (2008) found that teachers identified by students as engaging in differential treatment of high and low achievers were more likely to underestimate the ability of students from stereotyped ethnic groups. Further, teacher expectations contributed far more toward the year-end achievement gap (between stereotyped and nonstereotyped groups) in high bias classrooms ($d = .29$) than in low bias classrooms ($d = .003$). Similarly, Rubie-Davies (2007) found that in the classes of high expectation (for all) teachers, the mean effect size for student reading gains over one year was large ($d = 1.01$), while in the classes of low expectation teachers, the gains were very small ($d = .05$). This kind of variation in student exposure to teacher expectation effects can be likened to the variation documented in student access to quality of teachers across school years (NICHD Early Child Care Research Network, 2005).

Greater attention also needs to be paid toward how bias in teacher expectations is assessed and analyzed, over what school years, and with regard to what predictive domain, as differences in findings may result from different methods. While all of the studies of longer-term teacher expectation effects controlled for the prior performance of students, they varied in how teacher expectations were operationalized. Included were teachers' perceptions of intelligence (Alvidrez & Weinstein, 1999; Gut et al., 2013); of performance, talent, and effort (Smith et al., 1999); of reading and mathematics performance (Sorhagen, 2013) as well as of interest and engagement (Hinnant et al., 2009), recommended secondary school track placement (de Boer et al., 2010), and a prediction about whether or not youth would attend college (Mistry et al., 2007). It is perhaps timely to review the operationalization of expectations within the field and to work towards a more uniform measurement so that what is being measured is consistent across studies. Further, there may be critical time periods for teacher expectation influence (at the start of elementary or secondary schools) as well as critical student outcomes (such as the number of advanced mathematics courses a student might take) where the effects may be magnified.

Finally, future research might explore the practices of teachers, the structures of classroom, and the policies of schools that mediate between the expectations of teachers and their impact on student outcomes, especially as enacted across school years. Whereas there is substantial evidence for differential practices that heighten ability awareness and achievement gaps (e.g., Harris & Rosenthal, 1985; Weinstein, 2002) there is need to explore examples of cross-year transfer. As mentioned earlier, the Rist study (1970) described the fixed nature of reading group assignments, with its differential curricular exposure and labeling, as a potential mediating factor of cross-year effects.

Limitations

There are limitations to the study that should be noted. First, the study design was correlational, not experimental. Thus, analyses were tests of association, not of causal relations, even though the terms used in path analyses and CLPD use the term direct and indirect ‘effects.’ However, use of covariates (such as prior achievement or early verbal ability) and a longitudinal design (where earlier teacher expectations precede year-end student achievement), strengthen the claims about influence rather than prediction. It is also possible that teachers are more “accurate” than achievement scores in their expectations for students. They may be aware of other influences on student achievement and therefore teacher prescience, rather than influence, or omitted variables, may account for the resulting effects. Yet, having high expectations for student learning, beyond the level of current performance, has been found to be a positive facilitator of student growth. Research also documents that bias in teacher expectations, that is, over- and underestimation of ability relative to academic performance, is associated with differential treatment of students, of which students are well aware and may broaden or narrow the opportunity to learn (e.g., McKown & Weinstein, 2008). Further, these results are consistent with other studies that have shown long-term effects when teachers over- or underestimate (relative to achievement tests) the ability of students (e.g., Alvidrez & Weinstein, 1999; de Boer et al., 2010). In particular, teachers' underestimations of student ability for children whose homes were more educationally-oriented have been shown to be less predictive than for children whose homes had fewer educational resources, suggestive of a moderating variable of greater academic challenge from parents to which children responded (Alvidrez & Weinstein, 1999).

Second, the sample of students in the study was relatively small and homogeneous; generalizations from the study need to be considered with caution. The students were mostly European American, with high academic achievement, and their teachers had high expectations for them. Studies in the teacher expectation field generally suggest, for example, that expectation effects are greater for ethnic minority and low-income students than they are for white middle class students (Jussim et al., 1996; McKown & Weinstein, 2008; Rubie-Davies et al., 2006).

Conclusions

Research on the longer-term effects of teachers' expectations, that is, whether such effects endure, had prioritized the study of single teachers (e.g., Jussim & Harber, 2005; Smith et al., 1999) to the neglect of student histories of multiple teachers over successive school years. This study extends the evidence for the importance of teacher expectations across time to include direct and indirect pathways by which early teacher expectations predict later student achievement and the additive effects of multiple teacher expectation effects for a cohort of students. The study is strengthened by the inclusion of a preKindergarten measure of early verbal ability, achievement scores not seen by teachers, and a pure measure of teacher expectations focused on perceived intellectual ability. It does not confound perceptions of intellect with affective and social qualities as have some studies (e.g., Hinnant et al., 2009; Smith et al., 1999). Most importantly, this study shifts the debate from single teachers whose effects endure to the compounding of effects across multiple teachers over time — opening the door to contextual questions focused on students' experiences with different kinds of expectation environments across school years. The interesting next question is not whether long-term effects of teacher expectations exist on average but rather conditions for students and teachers under which expectation effects may be intensified and/or become more or less cumulative over the course of children's school careers (Weinstein, 2002).

This work heralds a new and contextual conception of longer-term effects of teacher expectations on student outcomes and paves the way for exciting research on longitudinal effects in the future. As yet, student experiences when they encounter teachers who have differential views of their capabilities have not been investigated. Further, given consistent findings in the literature about teacher qualities or practices that moderate teacher expectation effects, it is critical to pursue the role of teacher differences in the study of the longitudinal effects.

There is mounting evidence about classroom effects on children's achievement trajectories in early schooling (e.g., Pianta, Belsky, Vandergrift, Houts, & Morrison, 2008) yet we know relatively little about qualities of teachers and school environments that shape teachers' beliefs about what they can expect of students and how they actualize these beliefs in practice (Weinstein, 2008). While it is known that the prior achievement data teachers receive about students influences their expectations, it is also likely that within schools, the informal interactions of teachers also contribute to their expectations about students. Further, classrooms do not exist in isolation and teacher expectations can be influenced by school policies and preferences (Timperley & Robinson, 2001). Finally, future research should address the potentially different expectations of parents or other significant adults, which can compensate and interrupt negative achievement trajectories for children (e.g., Alvidrez & Weinstein, 1999).

An ecological or social contextual model could examine expectation processes as an interactive function of teacher and student differences and as nested within classrooms, grade levels, and schools as well as across time. To advance knowledge about expectation processes (both the positive and the negative) requires greater attention to moderator variables that define the conditions under which positive expectations flourish and thus contribute to enhanced and equitable learning opportunities for all students across their school careers.

Acknowledgments

The authors would like to acknowledge the assistance of Zena R. Mello, John A. C. Hattie and Gavin Brown for their early statistical support on this paper and Patricia Alexander for providing feedback on an earlier version of this manuscript.

References

Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first 2 years of school: Patterns and processes. *Monographs of the Society for Research in Child Development*, *53*, 1–157 (Serial No. 218).

Allison, P. (November 9, 2012). Why you probably need more imputations than you think. Retrieved from. http://www.statisticalhorizons.com/more-imputations

Alvidrez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology*, *91*, 731–746.

Babad, E. (2009). *The social psychology of the classroom.* New York, NY: Routledge.

Blatchford, P., Burke, J., Farquhar, C., Plewis, I., & Tizard, B. (1989). Teacher expectations in infant school: Associations with attainment and progress, curriculum coverage and classroom interaction. *British Journal of Educational Psychology*, *59*, 19–30.

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling*, *15*, 651–675.

Bohlmann, N. L., & Weinstein, R. S. (2013). Classroom context, teacher expectations, and cognitive level: Predicting children's math ability judgments, *34*, 288–298. http://dx.doi.org/10.1016/j.appdev.2013.06.003.

Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology*, *76*, 236–247.

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, *75*, 631–661.

Brophy, J. E. (1985). Teacher–student interaction. In J. B. Dusek (Ed.), *Teacher expectancies* (pp. 303–328). Hillsdale, NJ: Lawrence Erlbaum.

Brophy, J. E., & Good, T. L. (1974). *Teacher–student relationships: Causes and consequences.* New York, NY: Holt, Rinehart & Winston.

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context and reading: A comparison of students in 43 countries. *Scientific Studies of Reading*, *10*, 331–362.

Cowan, P. A., Cowan, C. P., Ablow, J. C., Johnson, V. K., & Measelle, J. R. (Eds.). (2005). *The family context of parenting in children's adaptation to elementary school.* Mahwah, NJ: Lawrence Erlbaum.

Cowan, C. P., Cowan, P. A., & Heming, G. (2005b). Two variations of a preventive intervention for couples: Effects on parents and children during the transition to school. In P. A. Cowan, C. P. Cowan, J. C. Ablow, V. K. Johnson, & J. R. Measelle (Eds.), *The family context of parenting in children's adaptation to elementary school* (pp. 277–312). Mahwah, NJ: Lawrence Erlbaum.

Cowan, P. A., Cowan, C. P., Heming, G., & Miller, N. (1995). *The child adaptive behavior inventory.* Berkeley, CA: Institute of Human Development, University of California at Berkeley.

Crano, W., & Mellon, P. (1978). Causal influence of teachers' expectations on children's academic performance: A cross-lagged panel analysis. *Journal of Educational Psychology*, *70*, 39–49.

de Boer, H., Bosker, R. J., & Van der Werf, M. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, *102*, 168–179. http://dx.doi.org/10.1037/a0017289.

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*, 1–17.

Donohue, K. M., Weinstein, R. S., Cowan, P. A., & Cowan, C. P. (2000). Patterns of teachers' whole-class perceptions and predictive relationships between teachers' and parents' perceptions of individual child competence. *Early Childhood Research Quarterly*, *15*, 279–305.

Enders, C., Dietz, S., Montague, M., & Dixon, J. (2006). Modern alternatives for dealing with missing data in special education research. *Advances in Learning and Behavioral Disabilities*, *19*, 101–129. http://dx.doi.org/10.1016/S0735-004X(06)19005-9.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*, 509–529.

Good, T. L., & Weinstein, R. S. (1986). Classroom expectations: One framework for exploring expectations. In K. K. Zumwalt (Ed.), *1986 ASCD yearbook: Improving teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Gut, J., Reimann, G., & Grob, A. (2013). A contextualized view on long-term predictors of academic performance. *Journal of Educational Psychology*, *105*, 436–443. http://dx.doi.org/10.1037/a0031503.

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*, 363–386.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London, UK: Routledge.

Hinnant, J. B., O'Brien, M. O., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology*, *101*, 662–670. http://dx.doi.org/10.1037/a0014306.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. http://dx.doi.org/10.1080/10705519909540118.

Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of preschool: Evidence from Virginia. *Early Childhood Research Quarterly*, *27*, 33–45. http://dx.doi.org/10.1016/j.ecresq.2011.03.006.

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. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, *Vol. 28*. (pp. 281–388). San Diego, CA: Academic Press.

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. http://dx.doi.org/10.1207/s15327957pspr0902_3.

Jussim, L., Robustelli, S. L., & Cain, T. R. (2009). Teacher expectations and self-fulfilling prophecies. In K. R. Wentzel, & A. Wigfield (Eds.), *Handbook of motivation in school* (pp. 349–380). New York, NY: Routledge.

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin*, *82*, 887–903.

Kenny, D. (2005). Cross-lagged panel design. In B. S. Everitt, & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, *Vol. 1*. (pp. 450–451). West Sussex, UK: Wiley.

Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kuklinski, M. R., & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development*, *72*, 1554–1578.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202.

Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *An introduction to statistical concepts.* New York, NY: Routledge.

Ma, X., & Xu, J. (2004). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education*, *110*, 256–280.

Marks, G. N. (2008). Are father's or mother's socioeconomic characteristics more important influences on student performance? Recent international evidence. *Social Indicators Research*, *85*, 293–309. http://dx.doi.org/10.1007/s11205-007-9132-4.

Markwardt, F. C. (1989). *Peabody Individual Achievement Test — Revised.* Circle Pines, MN: American Guidance Service.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction.* New York, NY: Guilford Press.

McKown, C., Gregory, A., & Weinstein, R. S. (2010). Expectations, stereotypes, and self-fulfilling prophecies in classroom and school life. In J. Meece, & J. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 256–274). Mahwah, NJ: Lawrence Erlbaum.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context and the achievement gap. *Journal of School Psychology*, *46*, 235–261. http://dx.doi.org/10.1016/j.jsp.2007.05.001.

Mistry, R. S., White, E. S., Benner, A. D., & Huynh, V. W. (2007). A longitudinal study of the simultaneous influence of mothers' and teachers' educational expectations on low-income youth's academic achievement. *Journal of Youth and Adolescence*, *38*, 826–838. http://dx.doi.org/10.1007/s10964-008-9300-0.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

NICHD Early Child Care Research Network (2005). A day in third grade: A large scale study of classroom quality and teacher and student behavior. *The Elementary School Journal*, *105*, 305–323.

Palardy, J. (1969). What teachers believe — What children achieve. *Elementary School Journal*, *69*, 370–374.

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, *45*, 365–397. http://dx.doi.org/10.3102/0002831207308230.

Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, *76*, 85–97.

Rist, R. C. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, *40*, 411–451.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258.

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

Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high and low expectation teachers. *British Journal of Educational Psychology*, *77*, 289–306. http://dx.doi.org/10.1348/000709906X101601.

Rubie-Davies, C. M. (2008). Teacher expectations. In T. Good (Ed.), *21st century education: A reference handbook* (pp. 254–262). Thousand Oaks, CA: Sage.

Rubie-Davies, C. M., Hattie, J., & Hamilton, R. (2006). Expecting the best for New Zealand students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology*, *76*, 429–444. http://dx.doi.org/10.1348/000709905X53589.

Rubie-Davies, C. M., Hattie, J. A., Townsend, M. A. R., & Hamilton, R. J. (2007). Aiming high: Teachers and their students. In V. N. Galwye (Ed.), *Progress in educational psychology research* (pp. 65–91). Hauppauge, NY: Nova.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys.* New York, NY: John Wiley.

Schafer, J. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, *8*, 3–15.

Schafer, J., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177. http://dx.doi.org/10.1037/1082-989X.7.2.147.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*, 323–338.

Smith, A. E., Jussim, L., & Eccles, J. (1999). Do self-fulfilling prophecies accumulate, dissipate, or remain stable over time? *Journal of Personality and Social Psychology*, *77*, 548–565.

Sobel, M. E. (1987). Direct and indirect effects in linear structural equation models. *Sociological Methods & Research*, *16*, 155–176. http://dx.doi.org/10.1177/0049124187016001006.

Sorhagen, N. S. (2013). Early teacher expectations disproportionately affect poor children's high school performance. *Journal of Educational Psychology*, *105*, 465–477. http://dx.doi.org/10.1037/a0031754.

Spitz, H. H. (1999). Beleaguered Pygmalion: A history of the controversy over claims that teacher raises intelligence. *Intelligence*, *27*, 199–234.

Stage, F. K., Carter, H. C., & Nora, A. (2004). Path analysis: An introduction and analysis of a decade of research. *The Journal of Educational Research*, *98*, 5–13. http://dx.doi.org/10.3200/JOER.98.1.5-13.

Timperley, H. S., & Robinson, V. M. J. (2001). Achieving school improvement through challenging and changing teachers' schema. *Journal of Educational Change*, *2*, 281–300.

Tizard, B., Blatchford, P., Burke, J., Farquhar, C., & Plewis, I. (1988). *Young children at school in the inner city.* Hove, UK: LEA.

Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling.* Cambridge, MA: Harvard University Press.

Weinstein, R. S. (2008). Schools that actualize high expectations for all youth: Theory for setting change and setting creation. In B. Shinn, & H. Yoshikawa (Eds.), *Toward positive youth development: Transforming schools and community programs.* London, UK: Oxford University Press.