THE
ECONOMIC
JOURNAL

The Economic Journal, **135** (*January*), 36–58 https://doi.org/10.1093/ej/ueae060
Advance Access Publication Date: 27 June 2024

DO MOTIVATED CLASSMATES MATTER FOR EDUCATIONAL SUCCESS?*

Jan Bietenbeck

I provide evidence of social spillovers of personality by showing that being in class with motivated peers affects educational success. I first document that academic motivation, a key aspect of personality in the context of education, predicts own achievement, classroom behaviour, the high school grade point average and college-test taking among elementary school students. Exploiting random assignment of students to classes, I then show that exposure to motivated classmates causally affects achievement, an effect that operates over and above spillovers of classmates' past achievement and socio-demographic composition. However, peer motivation in elementary school does not affect own motivation and long-term educational success.

A growing literature in economics and psychology documents the importance of personality for success in life (Borghans *et al.*, 2008; Almlund *et al.*, 2011; Heckman *et al.*, 2019). In particular, aspects of personality such as motivation, preferences and traits have been shown to predict performance in school and in the labour market (e.g., Duckworth *et al.*, 2007; Steinmayr and Spinath, 2009; Golsteyn *et al.*, 2014). Despite this crucial role played by personality in shaping individuals' own life outcomes, only very little research has examined how it affects other people in their social environment. This is surprising given that there is extensive evidence that peers matter for performance in school and in the workplace (e.g., Guryan *et al.*, 2009; Mas and Moretti, 2009; Sacerdote, 2011).

In this paper, I study the spillover effects of academic motivation, a key aspect of personality in the context of education. I use data from the Tennessee Student-Teacher Achievement Ratio experiment (Project STAR), which followed a single cohort of children from the beginning of kindergarten until the end of third grade. Two features make this setting uniquely suited for my purpose. First, the experiment measured students' academic motivation at the end of grades 1, 2 and 3 using a validated psychological scale. Second, some children entered the experiment in the second and third grades and were randomly assigned to existing classes within schools. This randomisation generated exogenous, observable variation in the predetermined motivation of entrants' classmates, which I can use to estimate causal spillover effects.

Psychologists define motivation as the conscious and unconscious needs and desires of individuals (Roberts, 2006; Roberts and Yoon, 2022). The scale used in Project STAR applies this definition to the context of learning and conceptualises academic motivation as consisting of two facets. First, *achievement needs* capture the utility that a child derives from learning and the associated social appreciation. Second, *failure avoidance* captures the disutility from low

* Corresponding address: Jan Bietenbeck, Department of Economics, Lund University, P.O. Box 7080, 220 07 Lund, Sweden. Email: mail@janbietenbeck.com

This paper was received on 27 September 2021 and accepted on 25 June 2024. The Editor was Marco Manacorda.

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.11911078.

Many thanks to Felix Weinhardt, Sule Alan, Fabian Kosse, Kaveh Majlesi, Therese Nilsson, Tove Nordström, Pia Pinger, Ulf Zölitz and numerous seminar and conference participants for helpful comments and suggestions. A previous version of this paper was entitled 'Peer Motivation and Educational Success'.

school achievement and the associated embarrassment. The scale measures these facets using a self-assessment questionnaire, with answers summarised in a single score for each student. This score captures children's academic motivation and possibly also other, related aspects of personality.

I begin my empirical analysis by showing that this motivation score predicts children's own educational success. I exploit the fact that participants were followed even after the experiment ended in order to study short- and long-term outcomes. The results reveal that, on average, children with a 1-standard-deviation (1-SD) higher motivation in elementary school score about 0.05 SDs higher on standardised reading and maths tests in elementary and middle school and are 4% more likely to take a college entrance exam around age eighteen. Motivation further predicts multiple measures of good classroom behaviour, as rated by teachers, in fourth and eighth grades.

I next investigate whether children's academic motivation affects the learning outcomes of their classmates. For this analysis, I focus on a sample of students who first entered Project STAR in the second or third grade. These students were randomly assigned to existing classes within school upon entry, which allows me to avoid the selection problems that typically complicate the identification of causal peer effects. Moreover, most new classmates of these entrants had participated in the experiment in the previous school year, which lets me observe their predetermined motivation. My regressions exploit the random variation in classmates' average motivation to identify spillover effects on entrants' short- and long-term educational success.

The results show that students who are randomly assigned to a class with more motivated peers initially perform better in school. Specifically, a 1-SD increase in classmates' average motivation raises performance on a standardised reading test at the end of the school year by 0.08 SDs (the effect on maths scores is 0.04 SDs, but this is imprecisely estimated). This spillover effect is not driven by an improvement in own motivation, which I show is unaffected by peer motivation. More generally, peer motivation does not seem to matter beyond contemporaneous achievement, as it does not affect any of the longer-term outcomes measured after the experiment ended and classes were reorganised at the end of third grade.

Peer motivation is likely correlated with other peer characteristics, which could potentially confound these estimates. In additional regressions, I show that controlling for classmates' past achievement and their composition in terms of gender, race and free-lunch eligibility does not change the estimated effect of peer motivation on test scores much. However, the data do not allow me to control for other aspects of peer personality. Therefore, the estimates are best understood as capturing the effects of peer motivation and other, related aspects of peer personality, which are distinct from spillovers due to peer ability and peer socio-demographic background.

What are the mechanisms behind these results? I argue that the spillovers on contemporaneous achievement are most likely due to an improved learning environment in school, as motivated peers show better classroom behaviour and distract their classmates less. As for the lack of longer-term effects, previous research has found that childhood interventions are particularly successful at changing future outcomes if they affect children's personalities (e.g., Heckman *et al.*, 2013). Thus, the absence of longer-term impacts might be due to the fact that peer motivation does not change own motivation. It appears that the contemporaneous effect on reading scores by itself is simply not large enough to generate measurable longer-term impacts. I briefly discuss the implications of these findings in the conclusion.

This paper contributes to a large literature on peer effects in education (for comprehensive surveys, see Sacerdote, 2011; Paloyo, 2020). One strand of this research focuses on spillovers from peer demographic composition, as measured, for example, by the share of female peers or the share of Black peers (e.g., Hoxby, 2000; Hoxby and Weingarth, 2005; Whitmore, 2005; Lavy and Schlosser, 2011; Brenoe and Zölitz, 2019). Another strand examines the consequences of being exposed to disruptive classmates (e.g., Figlio, 2007; Carrell and Hoekstra, 2010; Carrell *et al.*, 2018). Yet another strand studies spillovers from peer ability (e.g., Lavy *et al.*, 2012; Sojourner, 2013; Booij *et al.*, 2017; Feld and Zölitz, 2017). In Bietenbeck (2020), I add to this latter line of research by studying the impacts of being exposed to a very low-achieving repeater during kindergarten in Project STAR. While the current paper uses the same data, the treatment is very different, and I show below that the effects of peer motivation are robust to controlling for repeater exposure.

A few innovative recent papers extend the research on peer effects by studying spillovers from peer personality. Golsteyn *et al.* (2021) exploited data on personality traits and random assignment to classes in a university setting and found that students perform better in the presence of persistent peers, an effect that operates over and above spillovers from peer ability and peer demographic composition. Related work by Shure (2021) and Hancock and Hill (2022) showed that peer conscientiousness positively affects performance in early secondary school and in college, respectively. Ballis (2023) documented that peers of undocumented youths perform better in high school when the returns to schooling for these youths increase, an effect that could be driven by a boost in motivation among the undocumented students. I contribute to this research by studying spillovers from peer motivation in elementary school, when cognitive and non-cognitive skills are still highly malleable (Kautz *et al.*, 2014). Unlike previous studies, I can estimate effects on long-term outcomes. Moreover, I examine how peer motivation affects own motivation; together with parallel work by Shan and Zölitz (2022), this is the first evidence on whether peer personality affects own personality.

Finally, this paper also adds to the large literature in economics and psychology on the importance of personality (for surveys, see Borghans *et al.*, 2008; Almlund *et al.*, 2011; Heckman *et al.*, 2019). This research has shown that motivation (e.g., Wong and Csikszentmihalyi, 1991; Steinmayr and Spinath, 2009), preference parameters, such as patience (e.g., Golsteyn *et al.*, 2014; Cadena and Keys, 2015), and personality traits, such as conscientiousness (e.g., Poropat, 2009; Gensowski, 2018), grit (e.g., Duckworth *et al.*, 2007) and locus of control (e.g., Piatek and Pinger, 2016), predict educational success. Related recent work documents that school-based interventions can boost favourable aspects of personality in children and thereby improve their school performance (e.g., Alan and Ertac, 2018; Alan *et al.*, 2019; Sorrenti *et al.*, 2024). I complement this research by showing that academic motivation not only predicts children's own educational success, but also affects the learning outcomes of their peers.

1. Motivation in Personality Psychology

The prototypical model of personality in psychology conceives of a core of personality that is made up of four domains: traits, motives, abilities and narratives (Roberts, 2006; Roberts and Yoon, 2022). Traits capture the relatively stable patterns of thoughts, feelings and behaviours of an individual and are often represented using the well-known Big Five taxonomy. Motives are defined as what an individual desires, needs and strives for. Abilities capture things such as intelligence, and narratives are the stories that an individual tells herself in order to make sense of her life. How exactly these four domains relate to each other is the subject of an ongoing debate in psychology (Roberts and Yoon, 2022). However, it is widely accepted that together, they shape a person's identity and reputation, which in turn determine her roles in society.

This paper studies the importance of academic motivation, which falls under the motives domain. Unlike the literature on personality traits, psychological research on motivation has not converged on a common theoretical framework, system of measurement or terminology (Murphy and Alexander, 2000; Roberts *et al.*, 2006; Roberts and Yoon, 2022). Despite this heterogeneity, empirical studies have consistently found that motivation is predictive of success in life: for example, Steinmayr and Spinath (2009) documented that motivation predicts school performance over and above intelligence, and Dunifon and Duncan (1998) found that having an orientation toward challenge predicts future earnings. In related work in economics, Segal (2012) showed that intrinsic motivation in adolescence and early adulthood, as measured by performance on a low-stake coding speed test, predicts future earnings over and above cognitive skills.

The apparent importance of motivation for success in life has led psychologists to study potential ways to boost motivation among students. Results show that interventions that directly aim at increasing motivation, for example by helping students set learning goals or by instructing teachers to relate lesson content to students' experiences, can improve motivation and achievement (see Hulleman and Barron, 2015; Lazowski and Hulleman, 2016). In related research in economics, Heckman *et al.* (2013) showed that the Perry Preschool program boosted children's academic motivation, an effect that partly explains its positive impact on their longer-term educational success. In contrast, previous analyses of Project STAR did not find any evidence that class size affects motivation (Word *et al.*, 1990; Schanzenbach, 2006).

2. Project STAR: Background and Data

2.1. *Background on Project STAR*

Project STAR was a randomised controlled trial designed to investigate the effect of class size on student achievement. The original experiment followed a single cohort of children at seventy-nine schools in Tennessee from kindergarten through third grade. It started at the beginning of the 1985–6 school year, when 6,325 kindergarten students were randomly assigned to small classes (target size 13–17 students) or regular-sized classes (target size 22–5 students) within their school.[1] Because kindergarten was not mandatory at that time and due to normal residential mobility, 5,276 additional students joined this study cohort at participating schools during grades 1–3. These students were also randomised to classes within school upon entry, implying that the randomisation pool for all participants was school-by-entry-grade. After the initial randomisation, all students were supposed to stay in their assigned class type (small versus regular sized) until the end of third grade, at which point the experiment ended.

Teachers were also randomly assigned to classes within school at the start of each grade. As is common in the United States, Project STAR teachers worked in only one specific grade (that is, they were not 'looped'). As a consequence, students met a new, randomly assigned teacher in each and every grade.

As with any field experiment, the actual implementation of Project STAR deviated somewhat from the original plan. Thus, as children advanced from kindergarten to third grade, some students managed to move between small and regular-sized classes (for details, see

[1] There was also a third type of class: regular-sized class with a full-time teacher's aide. Previous studies using data from Project STAR have not found any differences in treatment effects between regular-sized classes with and without a full-time teacher's aide. In the empirical analysis, I follow the convention in the literature and group these two types of classes together.

Krueger, 1999). To account for this likely non-random sorting, I always define peer composition based on the initial random assignment when I estimate spillovers from motivated classmates below. Another deviation from the original study design was that a substantial number of students left the experiment either because they moved to other schools or because they were retained in grade. In Section 4 below, I provide evidence that this attrition is not driving my results.[2]

2.2. *Data and Variable Definitions*

My analysis is based on the Project STAR public use file (Achilles *et al.*, 2008), which allows me to follow students from the time they entered the experiment until the end of high school. In what follows, I give an overview of the main variables that I draw from this dataset. Additional details can be found in Online Appendix A.

2.2.1. *Academic motivation*

In the spring of each year from first through third grade, students' academic motivation was assessed using the early elementary form of the Self-Concept and Motivation Inventory (SCAMIN; Milchus *et al.*, 1968).[3] This psychological scale conceptualises academic motivation as consisting of two facets. First, *achievement needs* are defined as the positive regard with which a student perceives the intrinsic and extrinsic rewards of learning and performing in school. In economic terms, this captures the utility that a child derives from learning and the associated social appreciation.[4] Second, *failure avoidance* is defined as the awareness and concern toward shunning the embarrassment and sanctions that are associated with failure in school. In economic terms, this captures the disutility from low school achievement and the associated embarrassment.

As is common in personality psychology, the SCAMIN measures academic motivation using a self-assessment questionnaire. The instrument is group administered, which means that children complete the questionnaire in the classroom following instructions by their teacher. Specifically, students are first given an answer sheet containing twelve rows of five faces ranging from sad to happy. The teacher then reads out twelve corresponding questions starting with 'What face would you wear?' and asks students to mark the appropriate face as a response. For example, students are asked 'What face would you wear if you could read like a grown-up?' and 'What face would you wear if you could make your teacher happy with your arithmetic?' The questions cover both subject-specific achievement in reading and maths and school achievement in general. Half of the questions measure achievement needs and half measure failure avoidance.

The outcome of the assessment is a single academic motivation score for each student, which summarises her answers. In Project STAR, these scores were calculated centrally by the experimental staff following the SCAMIN scoring guidelines. These motivation scores, but not the

[2] For additional details on the design and implementation of Project STAR, see Word *et al.* (1990), Krueger (1999) and Finn *et al.* (2007).

[3] Motivation was also measured at the end of kindergarten using the preschool/kindergarten form of the SCAMIN. The questions and answer sheets of this form differ from those of the early elementary form, such that results are not comparable between the two forms. As described in detail in Online Appendix A, previous research in psychology and my own analyses using the Project STAR data cast serious doubt on the validity of the preschool/kindergarten form (but not the early elementary form). I therefore decided not to use kindergarten motivation in my analysis.

[4] I define this economic counterpart based on my analysis of the corresponding SCAMIN questions. Unfortunately, due to copyright restrictions, not all SCAMIN questions can be reproduced here, although two examples are given further below in the main text.

answers to individual questions, are included in the Project STAR public use file and form the basis for the empirical analysis below.

While the SCAMIN is designed to measure children's academic motivation, it could also capture other, related aspects of their personality. For example, the question about making the teacher happy mentioned in the previous paragraph might capture agreeableness, that is, how much the student is willing to cooperate with the teacher. Given this uncertainty, I choose to interpret my results conservatively as capturing the effects of academic motivation and other, related aspects of personality.

Finally, besides academic motivation, the SCAMIN also measures students' academic self-concept using a separate set of questions. Psychologists define self-concept as a person's perception of herself, which is formed through experience with her environment (Shavelson *et al.*, 1976). In the prototypical model of personality, self-concept forms part of a person's identity, which is shaped by the four core personality domains, but which may itself also influence these domains via feedback processes (Roberts, 2006). In Section 4 below, I study how peer motivation affects academic self-concept.

2.2.2. *Achievement in reading and maths*

At the end of each grade from kindergarten through third grade, participants in Project STAR wrote the grade-appropriate version of the Stanford Achievement Test. Moreover, in the spring of grades 5–8, all students who were enrolled in public schools in Tennessee wrote the Comprehensive Test of Basic Skills as part of a statewide testing program. Both tests are standardised assessments covering various subjects, and I use the reading and maths scores included in the Project STAR public use file as my main measures of student achievement.

2.2.3. *Classroom behaviour*

When STAR participants were in fourth grade, their teachers rated a subset of them on their classroom behaviour. Teacher ratings for twenty-eight behaviours were recorded on a scale from 1–5 and then consolidated into four indices. The effort index measures behaviours such as showing persistence when confronted with difficult problems. The initiative index captures things such as actively participating in classroom discussions. The discipline index measures behaviours such as being quiet versus interfering with classmates' work. The value index captures to what extent a student appreciates the school learning environment. All indices are coded such that higher values reflect better behaviour. In eighth grade, maths and English teachers rated a different subset of STAR participants using a similar, but shorter questionnaire, and the ratings were consolidated into the same four indices. In the analysis below, I measure classroom behaviour using the total of eight fourth- and eighth-grade indices.

2.2.4. *Educational attainment*

Most participants in Project STAR graduated from high school in 1998, and researchers collected information on the high school grade point average (GPA) and graduation status for participants attending selected high schools in 1999 and 2000. Besides this information, the public use file contains an indicator for whether a student had taken an ACT or SAT college-entrance test by 1998. This indicator is based on the administrative records of the two companies offering these tests and is the outcome of a data collection effort by Krueger and Whitmore (2001). It is available for the full sample of STAR participants and is a measure of college intent.

2.2.5. *Student characteristics*
The data contain information on the following socio-demographic characteristics of students: age, gender, race and an indicator for whether the student was ever eligible for free or reduced-price lunch during the experiment. Based on students' exact dates of birth and the school entry cut-off date in Tennessee, I additionally construct an old-for-grade indicator, which identifies students who either entered school late or repeated a grade. In my previous research on Project STAR, I found that old-for-grade students perform substantially worse in school compared to their on-grade peers (Bietenbeck, 2020). I also construct a measure of predicted achievement, which combines the socio-demographic characteristics such that they optimally predict students' reading and maths scores.

2.2.6. *Class size*
Most of my regressions control for the original Project STAR treatment: assignment to a small class. I measure treatment assignment upon entry into the experiment in order to avoid issues of non-compliance in later grades (see Section 2.1).

2.3. *Missing Data*

Like most other longitudinal data, the Project STAR data contain missing values in some variables, which could affect the results of my analysis. I distinguish between three cases of missing data. First, there are missing values in motivation scores. One main reason for this is a data matching problem: after teachers handed over the completed SCAMIN answer sheets to the experimental staff, many respondents could no longer be uniquely identified due to the lack of a consistently coded student identifier. If an answer sheet could not be uniquely matched, it was ignored, leading to missing motivation scores in the data (see Word *et al.*, 1990, p.210). Another important reason for missing values in the motivation variable is that many students only entered Project STAR in one of the later grades, and thus did not participate in the SCAMIN assessment in the earlier grades. The missing data imply that I do not usually observe the motivation of all students in a class, with the consequence that peer motivation is measured with error. In Section 4, I discuss this problem in detail and also provide solutions.

Second, there is missing information on some outcome variables for some students. The reasons are manifold and include purposeful selective data collection in order to save money and time (like with the classroom behaviour ratings and the high school outcomes), accidental selective data collection (for example, due to students being absent on the day of a test) and the loss of records (in particular, the lack of a unique student identifier meant that some test scores could not be matched to students; see Word *et al.*, 1990, p.209). A consequence of these missing outcome data for the empirical analysis is that sample sizes differ between regressions with different dependent variables. Importantly, I show in Section 4 below that peer motivation does not predict whether my main outcomes are observed for a given student, and that my results hold when the sample is restricted to students observed with all main outcomes.

Third, there are a few missing values in student socio-demographic variables, which I mostly use as controls in my regressions. In order not to reduce sample size unnecessarily, in all regressions in this paper I impute missing values in controls at the sample mean and include separate dummies for missing values on each control variable. Results are virtually identical if I instead exclude students with missing information on socio-demographic characteristics from the sample.

Table 1. *Correlates of Motivation.*

	Grade 1–3 motivation				
	(1)	(2)	(3)	(4)	(5)
Male	–0.292***				–0.285***
	(0.023)				(0.023)
Black		–0.026			–0.023
		(0.046)			(0.050)
Free lunch			–0.002		0.011
			(0.026)		(0.027)
Age in years				0.065*	0.078**
				(0.034)	(0.033)
Old for grade				–0.214***	–0.190***
				(0.047)	(0.047)
Small class	–0.000	–0.001	–0.001	–0.004	–0.003
	(0.027)	(0.028)	(0.028)	(0.028)	(0.027)
Observations	9,072	9,072	9,072	9,072	9,072

Notes: The table shows estimates of regressions of students' average motivation in grades 1–3 on student socio-demographic characteristics and a dummy for assignment to small class upon entry into Project STAR. The sample includes the 9,072 students for whom a motivation score is observed in at least one of grades 1, 2 and 3. The dependent variable is standardised to have mean 0 and SD 1. All regressions control for school-by-entry-grade fixed effects (regressions that omit these fixed effects yield very similar results). SEs reported in parentheses are clustered at the school-by-entry-grade level. * $p < .10$, ** $p < .05$, *** $p < .01$.

3. Academic Motivation: Correlates and Predictive Validity

3.1. *Sample Selection*

In this section, I examine how academic motivation correlates with students' socio-demographic characteristics and measures of their own contemporaneous and future educational success. For this descriptive analysis, I focus on the 9,072 Project STAR participants for whom I observe a motivation score in at least one of grades 1, 2 and 3.[5] I construct the average motivation of each student during these grades in three steps: (1) I standardise the motivation scores for each grade to have mean 0 and SD 1, (2) I average the available scores for each student across grades and (3) I standardise the resulting average scores to have mean 0 and SD 1. I prefer this measure of motivation because averaging across grades reduces measurement error and increases statistical precision by maximising sample size. Nevertheless, I also provide results for grade-specific measures of motivation, which are qualitatively similar.

3.2. *Correlates of Academic Motivation*

Table 1 shows estimates of regressions of average motivation in grades 1–3 on student socio-demographic characteristics, a small-class dummy and school-by-entry-grade fixed effects. Column (1) shows that male students are substantially less motivated on average, with a 0.29-SD lower motivation. In contrast, columns (2) and (3) show that there are no significant differences in motivation by race and free-lunch eligibility. Column (4) reveals that students who are old for grade are much less academically motivated, with a 0.21-SD lower motivation, and that, conditional on old-for-grade status, older students are slightly more motivated. Column (5) shows

[5] Of the students, 3,358 have a motivation score in only one grade, 2,401 have motivation scores in two grades and 3,313 have motivation scores in all three grades.

results from a regression in which all five student characteristics enter at the same time, which confirm the described patterns.

The final row in Table 1 shows the coefficients on the small-class dummy. Because assignment to small classes was random conditional on school-by-entry-grade fixed effects, these estimates capture the causal effect of class size on motivation. The results show that, unlike targeted interventions that directly aim to improve students' motivation (e.g., Hulleman and Barron, 2015; Lazowski and Hulleman, 2016), a non-targeted reduction in class size does not appear to boost students' motivation.

3.3. *Predictive Validity of Academic Motivation*

I now examine the predictive validity of academic motivation. I estimate regressions of the form

$$y_{is} = \alpha + \beta \mathrm{MOTIV}_i^{G1-G3} + \mathbf{X}_i'\gamma + \boldsymbol{\lambda}_s + \varepsilon_{is},$$

where i denotes students and s denotes school-by-entry-grade cells, that is, the Project STAR randomisation blocks; y_{is} is a measure of classroom behaviour or educational success; MOTIV_i^{G1-G3} is student i's average academic motivation across grades 1–3; $\mathbf{X}_i$ is a vector of socio-demographic controls; $\boldsymbol{\lambda}_s$ is a vector of school-by-entry-grade dummies, which account for differences between students entering the various schools participating in Project STAR in different grades. Finally, ε_{is} is the error term. In all regressions, I cluster SEs at the school-by-entry-grade level.

Table 2 reports the results. Panel A shows that motivation predicts good classroom behaviour, as rated by teachers, in fourth and eighth grades. For example, a 1-SD higher motivation in grades 1–3 is associated with 0.10-SD higher effort and 0.09-SD higher discipline in fourth grade. More motivated students also show better initiative and appreciate the school learning environment more. The associations are also positive, but slightly weaker for classroom behaviour in eighth grade, which could reflect either fade-out or the fact that the questions on which teachers rated students were different in that grade.

Panel B shows that in line with previous research from psychology (e.g., Wong and Csikszentmihalyi, 1991; Steinmayr and Spinath, 2009), motivation predicts short- and long-term educational success. For example, a 1-SD higher motivation is associated with 0.05-SD higher standardised reading and maths scores in both elementary school (grades 1–3) and middle school (grades 5–8). Motivation in early elementary school also predicts high school success and college intent: students with a 1-SD higher motivation have 0.3-point (0.04-SD) higher GPAs and are 1.5 percentage points more likely to take an ACT or SAT test around age eighteen, an increase that corresponds to about 4% of the sample mean.

How large are these associations? One way to gauge the size of the correlations between classroom behaviours and motivation is by comparing them to the gender gap, which has been widely documented in previous research (e.g., Bertrand and Pan, 2013). Across the eight measures of classroom behaviour studied in panel A of Table 2, the coefficients on motivation correspond to 22% of the gap between male and female students on average. Another salient reference point is the gap in educational outcomes between low- and high-socioeconomic-status students, as proxied by free-lunch eligibility. Panel B of Table 2 reveals that the estimated coefficients on motivation correspond to slightly more than 10% of the achievement gap in reading and maths between these two groups. Taken together, the results in Table 2 show that the motivation score

Table 2. *Own Motivation, Classroom Behaviour and Educational Success.*

Panel A: classroom behaviour

	Grade 4				Grade 8			
	Effort (1)	Initiative (2)	Discipline (3)	Value (4)	Effort (5)	Initiative (6)	Discipline (7)	Value (8)
Grade 1–3 motivation	0.102***	0.081***	0.090***	0.110***	0.060***	0.038*	0.065***	0.078***
	(0.027)	(0.026)	(0.028)	(0.031)	(0.022)	(0.023)	(0.024)	(0.024)
Male	–0.357***	–0.218***	–0.490***	–0.350***	–0.440***	–0.264***	–0.520***	–0.395***
	(0.041)	(0.044)	(0.049)	(0.042)	(0.039)	(0.046)	(0.041)	(0.043)
Free lunch	–0.423***	–0.472***	–0.228***	–0.248***	–0.304***	–0.266***	–0.242***	–0.198***
	(0.053)	(0.057)	(0.049)	(0.054)	(0.046)	(0.044)	(0.045)	(0.052)
Observations	2,212	2,212	2,212	2,212	2,693	2,693	2,693	2,693

Panel B: educational success

	Grades 1–3		Grades 5–8		High school		
	Reading (1)	Maths (2)	Reading (3)	Maths (4)	GPA (5)	Grad. (6)	ACT/SAT (7)
Grade 1–3 motivation	0.045***	0.051***	0.049***	0.052***	0.285*	0.007	0.015***
	(0.011)	(0.012)	(0.014)	(0.013)	(0.149)	(0.007)	(0.005)
Male	–0.184***	0.008	–0.100***	–0.138***	–3.080***	–0.070***	–0.133***
	(0.021)	(0.020)	(0.022)	(0.023)	(0.274)	(0.012)	(0.010)
Free lunch	–0.424***	–0.408***	–0.460***	–0.428***	–3.437***	–0.140***	–0.269***
	(0.029)	(0.028)	(0.028)	(0.028)	(0.374)	(0.014)	(0.015)
Observations	8,530	8,678	7,497	7,493	3,360	4,368	9,072

Notes: The table shows estimates from regressions of the outcome variables indicated in the column headers on students' academic motivation, averaged across grades 1–3. Measures of classroom behaviour in panel A are standardised to have mean 0 and SD 1. In columns (1)–(4) of panel B, test scores are averaged across the grades indicated in the column headers and are then standardised to have mean 0 and SD 1. The sample means of the high school outcomes used in columns (5)–(7) of panel B are 83.5 (GPA), 0.82 (graduation), 0.38 (ACT/SAT test taking). Sample sizes differ across outcomes because of different data collection procedures and sample attrition; see Online Appendix A for details. All regressions in panels A and B control for school-by-entry-grade fixed effects, dummies for male, Black, eligibility for free or reduced-price lunch, old for grade and age. SEs reported in parentheses are clustered at the school-by-entry-grade level. * $p < .10$, *** $p < .01$.

captures a dimension of personality that is reflected in students' actual behaviours and predictive of their educational success.[6]

4. Peer Motivation and Educational Success

4.1. *Sample Selection and Summary Statistics*

I now study how peer motivation affects educational success. Specifically, I estimate causal spillover effects on students who first entered Project STAR in the second or third grade. The new classmates of these entrants had participated in the experiment and written the SCAMIN test in the previous (first or second) grade, which allows me to observe their academic motivation. As students in Project STAR were randomly assigned to classes within school upon entry, this means that there is random and observable variation in the motivation of second- and third-grade entrants' classmates, which I can use to estimate causal spillover effects.

A total of 2,962 students entered Project STAR in the second or third grade. I construct peer motivation as the average motivation of these entrants' classmates at the end of the previous school year. This ensures that peer motivation is predetermined relative to the assignment of entrants to classes. For reasons described in Section 2, some classmates are not observed with a motivation score. In my main analysis, I ignore these missing values and compute peer motivation as the average of the available scores. Moreover, I drop ninety-four students from the sample for whom there is no information on any of their classmates' motivation. In Subsection 4.6 below, I describe the problem of missing peer motivation scores in more detail and discuss evidence showing that, as a consequence, my estimates are slightly biased toward zero.

For the remaining 2,868 students in the estimation sample, I construct a range of other peer variables, which I use as controls in some regressions. Specifically, I compute averages of classmates' socio-demographic characteristics and their reading and maths achievement in the previous grade. To facilitate interpretation and comparison of results, I standardise both peer achievement and peer motivation to have mean 0 and SD 1. I also construct a dummy for having a classmate who repeated kindergarten in the first year of Project STAR; this is the treatment I consider in Bietenbeck (2020), which captures exposure to a very low-achieving peer.

In line with the bulk of the previous research on peer effects, the main specifications focus on spillover effects on contemporaneous outcomes. Specifically, I estimate how exposure to motivated peers affects entrants' reading and maths achievement at the end of their first year in Project STAR. In additional analyses, I also examine impacts on entrants' own academic motivation and self-concept at the end of their first year in Project STAR and their long-term educational success. For ease of interpretation, I standardise all achievement outcomes to have mean 0 and SD 1.

Table 3 shows summary statistics for the peer motivation sample. Because of the fact that Project STAR oversampled schools in poor neighbourhoods, students are disproportionately likely to be Black and eligible for free lunch. Fully 47% of students are old for grade, which implies that they either entered school late or repeated a grade. In terms of outcomes, only 73%

[6] These findings are confirmed in regressions in which the main independent variable is grade-specific motivation, rather than average motivation across grades 1–3. Online Appendix Figure B1 reports the corresponding estimates, which are based on a consistent sample of students observed with motivation scores in all three grades. While motivation in each grade is positively associated with most outcomes, the point estimates tend to be smaller than those for average (across grades) motivation. This supports the intuition that averaging reduces measurement error in the motivation variable.

Table 3. *Summary Statistics for the Peer Motivation Sample.*

	Mean	SD	*N*
Socio-demographic characteristics			
Male	0.55	0.50	2,861
Black	0.42	0.49	2,766
Free lunch	0.66	0.47	2,730
Age in 1985	6.01	0.70	2,845
Old for grade	0.47	0.50	2,845
Peer motivation and other peer characteristics			
Peer motivation	0.00	1.00	2,868
Peer reading achievement	0.00	1.00	2,841
Peer maths achievement	0.00	1.00	2,850
KG repeater peer in class	0.21	0.40	2,868
Peer share male	0.51	0.11	2,868
Peer share Black	0.42	0.43	2,868
Peer share free lunch	0.61	0.30	2,868
Entry-grade achievement			
Reading score	0.00	1.00	2,185
Maths score	0.00	1.00	2,196
Entry-grade own personality			
Own motivation	0.00	1.00	2,276
Own self-concept	0.00	1.00	2,276
Long-term educational outcomes			
Reading scores in grades 5–8	0.00	1.00	2,118
Maths scores in grades 5–8	0.00	1.00	2,119
High school GPA (0–100)	81.50	7.46	665
High school graduation	0.73	0.44	1,018
Took ACT/SAT	0.26	0.44	2,868

Notes: The table shows means, SDs and the number of students observed with each variable for the 2,868 students included in the peer motivation sample. KG repeater refers to a child who repeated kindergarten in the first year of Project STAR.

of students graduated from high school and only 26% took an ACT or SAT test around the age of eighteen. Taken together, these statistics show that the sample mostly includes disadvantaged and low-achieving students.

4.2. *Regression Specification*

I estimate regressions of the form

$$y_{ics} = \theta \overline{\mathrm{MOTIV}}_c^{G-1} + \phi \mathrm{SMALL}_c + \mathbf{X}_i' \eta + \overline{\mathbf{Z}}_c' \rho + \boldsymbol{\omega}_s + \mu_{ics}, \tag{1}$$

where i denotes students, c denotes classes and s denotes school-by-entry grade cells; y_{ics} is the outcome of interest; $\overline{\mathrm{MOTIV}}_c^{G-1}$ is the average motivation of students in class c who participated in Project STAR in the previous grade ($G-1$)—as described above, this average is computed based only on the non-missing motivation scores; SMALL_c is a dummy for assignment to a small class, the original treatment of interest in Project STAR; $\mathbf{X}_i$ is a vector of student socio-demographic characteristics and $\overline{\mathbf{Z}}_c$ is a vector of predetermined peer characteristics shown in Table 3. Finally, $\boldsymbol{\omega}_s$ is a vector of school-by-entry-grade dummies that accounts for fixed differences between

Table 4. *Balancing Tests for Peer Motivation and Peer Achievement.*

	Male (1)	Black (2)	Free lunch (3)	Age (4)	Old for grade (5)	Pred. achievement (6)
Panel A: separate regressions for each peer variable						
Peer motivation	0.002	–0.007	–0.005	–0.023	–0.004	0.014
	(0.012)	(0.006)	(0.009)	(0.017)	(0.011)	(0.018)
Peer reading achievement	0.017	–0.008	–0.014	–0.024	–0.005	0.024
	(0.015)	(0.009)	(0.021)	(0.023)	(0.015)	(0.029)
Peer maths achievement	0.024	–0.012	–0.028*	–0.020	–0.010	0.038
	(0.015)	(0.010)	(0.016)	(0.028)	(0.019)	(0.031)
Panel B: joint regressions for all peer variables						
Peer motivation	0.002	–0.007	–0.004	–0.022	–0.004	0.014
	(0.012)	(0.006)	(0.010)	(0.016)	(0.011)	(0.018)
Peer reading achievement	–0.000	0.002	0.009	–0.015	0.003	–0.006
	(0.021)	(0.010)	(0.029)	(0.029)	(0.019)	(0.038)
Peer maths achievement	0.024	–0.013	–0.033	–0.009	–0.012	0.042
	(0.021)	(0.012)	(0.020)	(0.036)	(0.025)	(0.039)
p-value (joint significance)	0.44	0.37	0.22	0.42	0.95	0.59
Observations (both panels)	2,861	2,766	2,730	2,845	2,845	2,868

Notes: The table shows estimates of regressions of students' socio-demographic characteristics and predicted achievement on the characteristics of their classmates. Estimates are based on the peer motivation sample. In panel A, each coefficient comes from a separate regression of the outcome indicated in the column header on the peer variable indicated in the row. In panel B, coefficients are instead based on a single regression in which all peer variables enter jointly. The *p*-value reported in panel B comes from an *F*-test for the joint significance of the three peer variables. All regressions in both panels control for school-by-entry-grade fixed effects. SEs reported in parentheses are clustered at the school-by-entry-grade level. * $p < .10$.

randomisation pools and μ_{ics} is the error term. For all regressions, I compute SEs that allow for clustering at the school-by-entry-grade level.

Equation (1) corresponds to a linear-in-means model, which is the most widely estimated model of peer effects (Sacerdote, 2011). The main coefficient of interest, θ, captures the causal impact of exposure to motivated peers under the assumption that variation in peer motivation is random within school-by-entry-grade cells, an assumption that I support with empirical evidence below. Since peer motivation is correlated with other peer characteristics, an obvious question is whether θ captures spillovers from motivation or from such other characteristics. I address this question by controlling for peers' previous achievement and socio-demographic characteristics, the main variables used to study peer effects in the previous literature (see Sacerdote, 2011; Paloyo, 2020). If estimates are robust to the inclusion of these controls, this suggests that θ indeed captures spillovers from peer motivation (and other, correlated aspects of personality not captured by these controls; see Altonji *et al.*, 2005; Oster, 2019).

4.3. *Evidence on Random Assignment*

I now provide evidence that students were indeed randomly assigned to classes within school upon entry. Specifically, I show that peer motivation is unrelated to predetermined characteristics of students entering the experiment in the second or third grade. Table 4 reports results from regressions like in (1) in which the dependent variables are students' predetermined socio-demographic characteristics (columns (1)–(5)) and predicted achievement (column (6)). Panel A shows estimates from separate regressions for peer motivation and, to further buttress the results, peers' past achievement in reading and maths. Panel B shows estimates from specifications in

Table 5. *Peer Motivation and Entry-Grade Achievement.*

	Reading			Maths		
	(1)	(2)	(3)	(4)	(5)	(6)
Peer motivation	0.081***	0.074***	0.071***	0.036	0.032	0.027
	(0.023)	(0.023)	(0.024)	(0.032)	(0.031)	(0.032)
Peer reading achievement		0.154**	0.152**		0.150**	0.134**
		(0.064)	(0.066)		(0.067)	(0.067)
Peer maths achievement		0.038	0.042		0.051	0.062
		(0.058)	(0.059)		(0.057)	(0.058)
KG repeater peer in class		–0.069	–0.077		0.004	–0.003
		(0.073)	(0.073)		(0.086)	(0.088)
Peer share male			–0.194			–0.421*
			(0.271)			(0.233)
Peer share free lunch			0.146			0.006
			(0.252)			(0.282)
Peer share Black			0.158			0.036
			(0.307)			(0.333)
Observations	2,185	2,185	2,185	2,196	2,196	2,196

Notes: The table shows estimates of the effect of peer motivation on achievement in reading (columns (1)–(3)) and maths (columns (4)–(6)) at the end of students' first year in Project STAR. Estimates are based on the peer motivation sample. All regressions control for own socio-demographic characteristics, a dummy for small class and school-by-entry-grade fixed effects. Regressions in columns (2), (3), (5) and (6) additionally control for averages of classmates' reading and maths achievement in the previous school year and an indicator for whether the class includes a kindergarten repeater, and regressions in columns (3) and (6) additionally control for averages of classmates' socio-demographic characteristics. SEs reported in parentheses are clustered at the school-by-entry-grade level. * $p < .10$, ** $p < .05$, *** $p < .01$.

which these three peer variables enter simultaneously instead. Across all regressions, most of the coefficients on the peer variables are close to zero and not statistically significant at conventional levels. In the regressions in panel B, the coefficients are also jointly insignificant.

In Online Appendix B, I present and discuss evidence from three further balancing tests. Like the estimates in Table 4, the outcomes of these tests strongly suggest that students were indeed randomly assigned to classes within school, supporting the validity of my empirical approach.

4.4. *Main Results: Effects on Contemporaneous Achievement*

Table 5 reports my main estimates of the effect of exposure to motivated peers on reading and maths achievement at the end of entrants' first year in Project STAR. Column (1) shows that having classmates with a 1-SD higher average motivation raises own reading achievement by 0.081 SDs. Column (4) shows an effect on maths achievement that is also positive, but smaller at 0.036 SDs and not statistically significant at conventional levels. Figure 1 visualises these estimates and reveals that the effects are roughly linear in average peer motivation.

Columns (2) and (5) of Table 5 add three controls for peer ability to these regressions: classmates' average reading and maths achievement in the previous school year and an indicator for whether the class includes a very low-achieving kindergarten repeater. If spillovers from motivated peers were mainly due to correlated peer ability, we would expect this to lead to a substantial reduction in the size of the coefficient on peer motivation. However, the estimates are largely unchanged, suggesting that this is not the case. Columns (3) and (6) show that the results are also robust to controlling for classmates' socio-demographic characteristics. This suggests that the coefficient on peer motivation captures a true personality spillover. However, as noted before, I am unable to control for other aspects of personality that might be correlated with

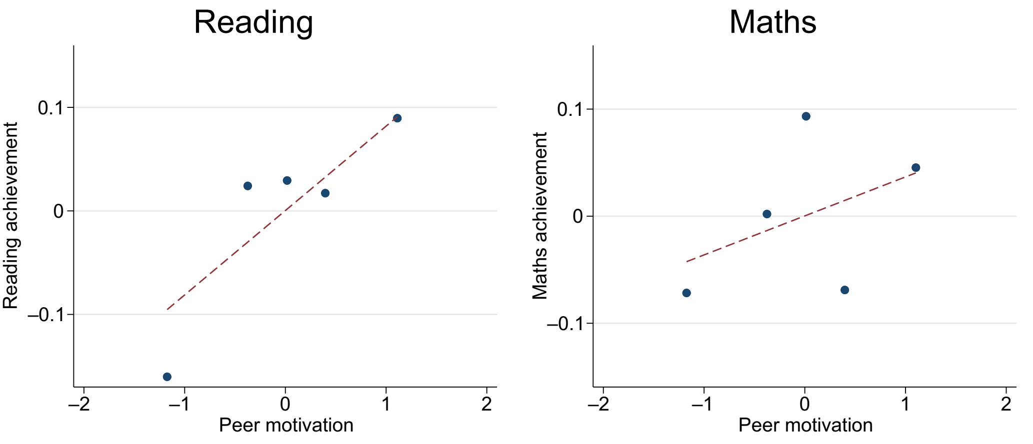

Fig. 1. *Peer Motivation and Entry-Grade Achievement.*
Notes: The figure shows estimates of the effect of peer motivation on reading and maths achievement at the end of entrants' first year in Project STAR. To construct these plots, I first residualise achievement scores and peer motivation on the controls included in the specifications in columns (1) and (4) of Table 5. I then group residualised peer motivation into ten equal-sized bins and plot the mean of the residualised achievement scores for each bin. The regression line in each plot is based on the underlying individual-level data and thus visualises the corresponding regression in Table 5.

motivation.[7] Therefore, the estimates in Table 5 are best interpreted as capturing the effects of peer motivation and other, correlated aspects of peer personality.

In additional analyses, I explore whether the effect of peer motivation differs by entrants' socio-demographic characteristics and two widely studied educational inputs, class size and teacher experience. Figure 2 presents results from regressions in which the sample is split into corresponding subgroups. The effect appears to be larger for boys, Black students and on-grade students, although none of these differences is statistically significant at conventional levels.[8] It also appears to be larger (though not significantly so) in regular-sized classes as compared to small classes, and in classes with less experienced teachers. Overall, while these analyses point toward potential heterogeneities in the effect of peer motivation, the relatively small sample size means that I lack statistical power to draw definitive conclusions.

4.5. *Further Results: Effects on Own Motivation and Long-Term Outcomes*

I now examine the effects of peer motivation on further outcomes. First, an intriguing possibility is that peer personality affects own personality. In columns (1) and (2) of Table 6, I explore such spillovers by estimating the effect of peer motivation on entrants' own motivation and self-

[7] The one exception is peer self-concept. Online Appendix Table B4 shows that the impacts of peer motivation are robust to controlling for peer self-concept in the regressions.

[8] I also tested whether there are distinct spillover effects from male versus female classmates' motivation. Online Appendix Table B5 shows that this is not the case in general. However, the estimates do suggest that male students benefit disproportionately from having motivated male peers, and female students benefit disproportionately from having motivated female peers, although these differences are not statistically significant at conventional levels. Furthermore, I examined whether the effects of peer motivation differ by length of exposure; however, the relatively small sample size and the complicating fact that students enter and exit the experiment in each grade did not allow me to draw firm conclusions about this potential heterogeneity.

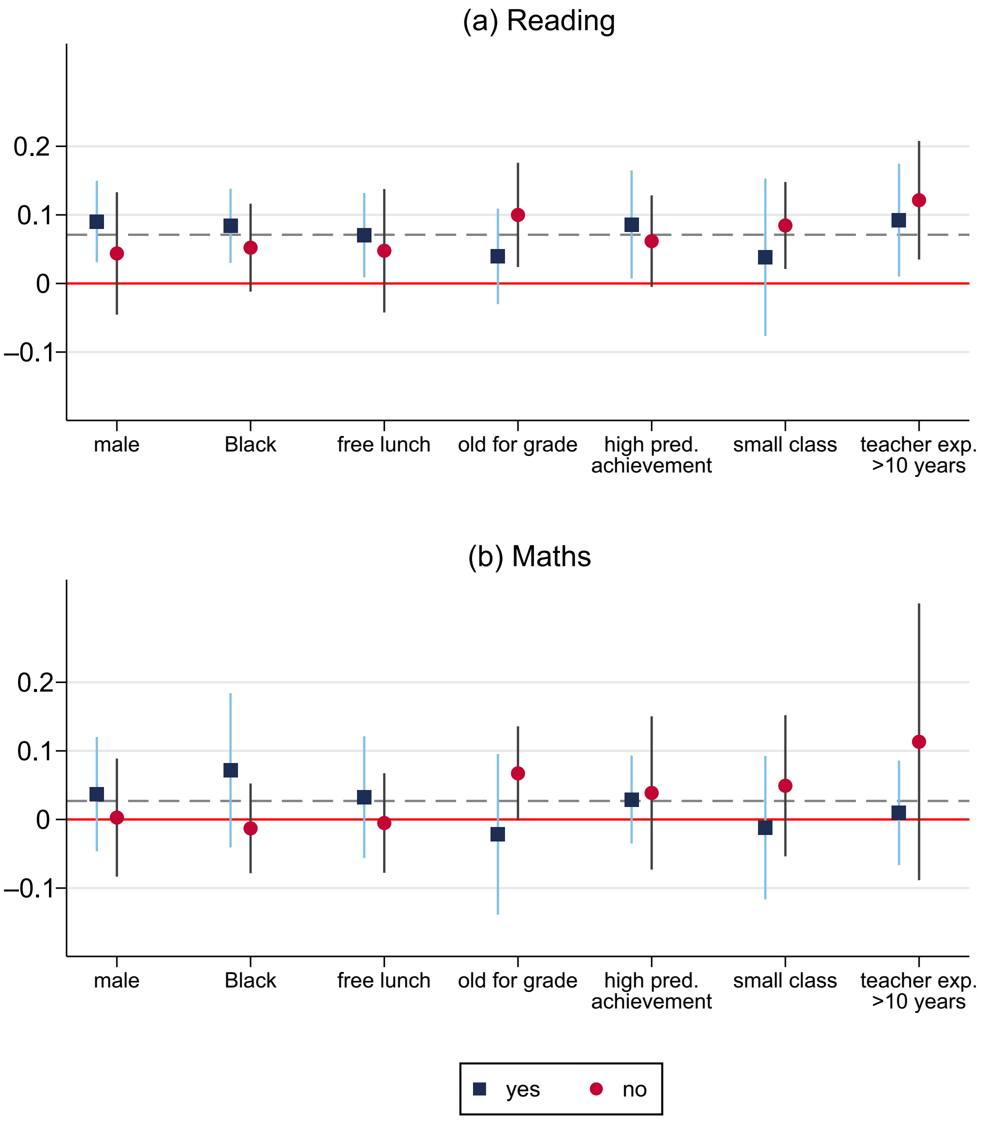

Fig. 2. *Peer Motivation and Entry-Grade Achievement: Heterogeneity Analysis.*
Notes: The figure shows point estimates and 95% confidence intervals of the effect of peer motivation on achievement in reading (panel (a)) and maths (panel (b)), separately for different groups of students. The specifications correspond to those in columns (3) and (6) of Table 5, but focus on subsamples of students as indicated on the horizontal axes: squares indicate point estimates for students with the respective characteristic, and circles indicate point estimates for students without this characteristic. High predicted achievement is an indicator for whether predicted achievement is above average. The ten-year cutoff for teacher experience is chosen for consistency with the analysis of the role of teacher experience in Project STAR in Chetty *et al.* (2011). The dashed line in each panel shows the main estimate for all students from columns (3) and (6) of Table 5.

Table 6. *Peer Motivation, Entry-Grade Own Motivation and Self-Concept, and Long-Term Educational Success.*

	Entry grade		Grades 5–8		High school		College
	Motivation (1)	Self-concept (2)	Reading (3)	Maths (4)	GPA (5)	Grad. (6)	ACT/SAT (7)
Peer motivation	–0.004	0.000	–0.024	–0.026	–0.467	–0.030*	–0.007
	(0.028)	(0.028)	(0.020)	(0.022)	(0.419)	(0.017)	(0.009)
Peer achievement controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,276	2,276	2,118	2,119	665	1,018	2,868

Notes: The table shows estimates of the effect of peer motivation on the outcome variables indicated in the column headers. Regressions control for own socio-demographic characteristics, averages of classmates' reading and maths achievement in the previous school year, an indicator for whether the class includes a kindergarten repeater, averages of classmates' socio-demographic characteristics, a dummy for small class and school-by-entry-grade fixed effects. SEs reported in parentheses are clustered at the school-by-entry-grade level. * $p < .10$.

concept at the end of their first year in Project STAR. The estimated effect of peer motivation in both regressions is almost exactly zero, showing that peer motivation does not affect own motivation or self-concept.

Second, given that peer motivation raises contemporaneous achievement, an obvious question is whether it also affects students' long-term educational success. I address this question by estimating effects on middle school test scores, high school outcomes and college-test taking. When interpreting these estimates, it is important to realise that they capture the impacts of a relatively short exposure to more motivated peers during early elementary school. Specifically, when Project STAR ended after third grade, students were redistributed to ordinary classes. While I do not observe class composition beyond third grade, this re-shuffling likely meant that peer motivation in the second or third grade was at most weakly related to peer motivation in later grades. Therefore, my estimates reflect the effects of differential exposure to more motivated peers for only one or two years during early elementary school.

Columns (3) to (7) of Table 6 show the results from this long-term analysis. Across the five regressions, there is no indication that the short-term positive spillover from motivated peers on achievement translates into longer-term educational success. If anything, the estimates point toward a negative effect of peer motivation on later outcomes, although most coefficients are imprecisely estimated and I cannot exclude small positive effects. I discuss potential explanations for this apparent discrepancy between short- and long-term effects of peer motivation in Section 5 below.

4.6. *Robustness*

In what follows, I summarise the results of robustness checks that address potential concerns about the validity of my findings. Online Appendix B provides a more detailed discussion of these analyses.

4.6.1. *Missing data on peer motivation*

As described in Section 2.3, motivation is not observed for all students, implying that peer motivation is measured with error. On average, 67% of classmates have motivation scores. This lack of data is mostly attributable to new entrants: if several students enter a given class in the same grade, the co-entrants of any given entrant mechanically do not have motivation scores because they did not participate in Project STAR in the previous year. When such co-entrants are excluded, the share of classmates observed with motivation scores rises to 86%.

Under random assignment to classes, missing information on classmates' motivation attenuates estimates toward zero if the peer average is constructed only from the available information. This implies that my main results underestimate the effects of peer motivation. Using a correction developed by Sojourner (2013) in the context of Project STAR, I find that this bias is relatively small: the estimated effect on entry-grade reading achievement rises from 0.071 to 0.096 SDs, but none of the estimates for the other outcomes turns statistically significant at conventional levels.

4.6.2. *Missing outcome data*

For reasons detailed in Section 2.3, not all outcomes are observed for all students in the sample. This opens up the possibility that my results are biased by selective attrition. I provide two pieces of evidence that this is not the case. First, I show that peer motivation does not predict whether outcomes are observed for a given entrant. Second, I document that restricting the sample to students observed with all main outcomes yields estimates that are very similar to my main results. Taken together, these analyses suggest that missing outcome data do not bias my results.

4.6.3. *Multiple hypothesis testing*

I study effects on many different outcomes, which raises the possibility that the only statistically significant effect on contemporaneous reading achievement represents a chance finding. I address this concern by showing that peer motivation also affects word study skills, which are closely related to reading skills and which were also assessed by the Stanford Achievement Test. Moreover, I show that the effects of peer motivation on reading and word study skills remain statistically significant when I correct for multiple hypothesis testing using the method developed by Romano and Wolf (2005a,b). These results suggest that the effects of peer motivation on contemporaneous achievement are unlikely to be a mere chance finding.

5. Discussion, Mechanisms and Policy Implications

5.1. *Discussion of Main Results*

The results in Section 4 show that exposure to motivated peers in early elementary school increases achievement on standardised tests. How does the size of these short-term spillovers compare with that of other estimates of peer effects in education? Table 5 shows that the effect on reading achievement is about half as large as the effect of a 1-SD increase in peers' past reading achievement and about the same size, in absolute value, as the effect of being exposed to a kindergarten repeater in the same sample.[9] With respect to the few existing estimates of spillovers from peer personality, my estimates for reading are larger than those found in higher education settings by Golsteyn *et al.* (2021) and Hancock and Hill (2022), whose main spillover estimates are 0.02 and 0.03 SDs, respectively, but smaller than those found for twelve-year-old students by Shure (2021), whose main spillover estimates range from 0.12 to 0.15 SDs. While

[9] In Bietenbeck (2020), I also document a negative effect of exposure to kindergarten repeaters on contemporaneous achievement. However, that paper focuses on first-time kindergarten students as the treated group, a sample that is significantly less disadvantaged than the second- and third-grade entrants considered here. In contrast to the results shown in Table 5, the effect of repeater exposure on kindergarten achievement is much larger for maths than for reading.

comparing estimates across different dimensions of personality and settings is difficult, these results appear broadly in line with the idea that skills are more malleable early in life and that therefore spillovers from peer personality are stronger at earlier ages.

Turning to long-term effects, throughout my analyses, I find no evidence that peer personality in early elementary school affects educational outcomes beyond the short term. This is somewhat surprising: given dynamic complementarities, one would expect some longer-term effects. However, it is important to note that the pattern of impacts is consistent with previous studies on childhood interventions, which have found that treatments are particularly successful at changing longer-term outcomes if they affect children's personality (e.g., Heckman *et al.*, 2013), and with earlier papers on peer effects, which have argued that school peers influence children's long-term educational and labour market success mainly via their impact on non-cognitive skills (e.g., Carrell *et al.*, 2018; Bietenbeck, 2020). Thus, the absence of longer-term impacts of peer motivation might be due to the lack of an effect on own motivation. Perhaps the contemporaneous impact on reading scores by itself is simply not large enough to generate measurable long-term effects.

In the end, however, I cannot provide definite evidence on why the short-term and long-term impacts appear to differ.[10] Factors that cannot be observed in the data, such as compensatory behaviour by parents, might play a role. Moreover, the precision of my estimates does not let me rule out small positive effects of peer motivation on long-term outcomes also. Ultimately, the question of whether peer personality also matters for long-term educational success will therefore have to be answered by future research.

5.2. *Mechanism for Short-Term Spillovers*

I now discuss potential mechanisms behind the effect of peer motivation on contemporaneous achievement. First, the experimental setup lets me rule out the most obvious explanations that involve selection into peer groups, sorting to specific teachers and selection into the sample. Second, another intuitive explanation is that peer motivation influences children's own personality, which in turn affects achievement. However, my results above provide no evidence of such personality change. A related possibility is that exposure to motivated peers changes students' norms about studying or doing homework. While I cannot observe such norms, this explanation is difficult to reconcile with the null effect on own motivation and with the apparent differences in effect size by class size and teacher experience, as it is unclear ex ante why studying norms should be influenced by these variables.

Third, yet another alternative mechanism is that motivated peers create a good learning environment in the classroom. As shown in Section 3, motivated students score higher on the discipline index, which measures the extent to which they (do not) interfere with their classmates' learning. Motivated students are also rated higher on other dimensions of good classroom behaviour by their teachers. This implies that entrants whose peers are more motivated likely experience less distraction from them, which in turn could account for the documented increase in achievement. While I cannot provide direct evidence in favour of it, I consider this the most likely mechanism behind the positive spillover effects from motivated peers.

[10] The same goes for why the effect of peer motivation differs between reading and maths, despite the fact that the SCAMIN measures motivation related to both subjects and that the associations of motivation scores with achievement in both subjects are virtually identical (see Section 3).

Table 7. *Peer Motivation and Entry-Grade Achievement: Bad Apples and Shining Lights.*

	Reading			Maths		
	All students	By pred. achievement		All students	By pred. achievement	
		Low	High		Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Share of peers with top 33% motivation	0.136	0.054	0.432	0.074	0.250	0.066
	(0.187)	(0.285)	(0.289)	(0.295)	(0.475)	(0.315)
Share of peers with bottom 33% motivation	–0.429***	–0.459**	–0.310	–0.222	–0.290	–0.109
	(0.157)	(0.218)	(0.278)	(0.174)	(0.285)	(0.244)
Peer achievement controls	Yes	Yes	Yes	Yes	Yes	Yes
Peer demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,185	1,143	1,042	2,196	1,142	1,054

Notes: The table shows estimates of the effect of peer motivation on achievement in reading and maths. Peer motivation is measured as the shares of classmates with top 33% and bottom 33% motivation scores. Regressions control for own socio-demographic characteristics, averages of classmates' reading and maths achievement in the previous school year, an indicator for whether the class includes a kindergarten repeater, averages of classmates' socio-demographic characteristics, a dummy for small class and school-by-entry-grade fixed effects. SEs reported in parentheses are clustered at the school-by-entry-grade level. ** $p < .05$, *** $p < .01$.

5.3. *Policy Implications*

Focusing on the short-term impact of peer motivation on achievement, I now consider potential policy implications of my findings. The often implicit promise of peer effects is that one may be able to improve average student outcomes by optimally assigning students to classes. Importantly, any such improvement in average outcomes requires peer effects not to be linear in means. I test for such non-linearity by asking whether exposure to peers with particularly low or particularly high motivation has a disproportionate effect on achievement, in line with the 'bad apple' and 'shining light' models of peer effects suggested by Hoxby and Weingarth (2005). For this purpose, I replace the average peer motivation term in (1) with the shares of classmates with top-tercile and bottom-tercile motivation scores. I estimate effects both for the full sample of students and separately for students with low and high predicted achievement.

Table 7 shows the results. Columns (1) and (4) reveal that the effect of peer motivation is driven by students with very low motivation, as exposure to such 'bad apples' has a large negative effect on achievement. The results in the other columns further reveal that low-predicted-achievement entrants are disproportionately hurt by the presence of such students. One potential implication of these results is that average achievement might be improved by systematically assigning low-predicted-achievement students to more motivated peers. Alternatively, students with very low motivation might be placed into separate classes (although my data do not allow me to estimate how such students affect each other). However, Carrell *et al.* (2013) provided a cautionary tale of actually implementing such 'optimal' assignment policies: they showed that endogenous peer group formation may offset the predicted gains from reassignment, something that I cannot rule out would happen even in my setting.

Rather than providing a blueprint for optimally assigning students to classes, my results speak to the kind of targeted programs that previous research has shown can effectively change aspects of personality, including motivation, in children. In particular, my findings suggest that the benefits of such interventions may be underestimated, as the generated improvements in personality for treated children will positively affect the learning outcomes of their peers. Incorporating such spillover benefits in the evaluation of such interventions thus appears important.

Building on the prior discussion, an important question emerges regarding which aspects of personality matter most for individual and peer educational success, as well as the type of personality data schools ought to gather to inform policy. My analysis underscores the significance of academic motivation, while other studies highlight the relevance of peer personality traits such as persistence (Golsteyn *et al.*, 2021), Big Five traits (Shure, 2021; Hancock and Hill, 2022; Shan and Zölitz, 2022) and competitiveness (Shan and Zölitz, 2022). Unfortunately, the very different contexts across these studies complicate direct comparison. Consequently, there is a need for additional research to refine our understanding of the interplay between different aspects of personality and educational success.

6. Conclusion

Previous research in economics and psychology has documented the importance of personality for individuals' own life success. However, despite extensive evidence that peers matter for performance in school and in the workplace, only very few studies have examined spillovers of personality in the social environment. This paper helps fill this gap by showing that academic motivation, which is a key aspect of personality in the context of education, affects peers' educational success.

My empirical analysis exploits the random assignment of students to classes in elementary schools in Project STAR. I find that being assigned to more motivated classmates causally increases achievement on a standardised reading test at the end of the school year. This peer effect operates over and above spillovers of classmates' academic ability and socio-demographic composition, which suggests that it reflects a true personality spillover. Since peer motivation does not affect own motivation, I argue that the positive spillover on achievement is most likely due to an improved classroom learning environment: as I show, motivated students tend to distract their classmates less. The lack of an effect on own motivation also offers an explanation for the null effect of peer motivation on longer-term educational success.

My findings suggest that the benefits of interventions that positively affect children's personalities may be underestimated, as the generated improvements for treated children will positively affect the learning outcomes of their peers. More generally, I show that the effects of any educational input that has an impact on personality may extend beyond the students who are targeted, as personality affects other people in their social environment.

Lund University, Sweden

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Achilles, C., Bain, H.P., Bellott, F., Boyd-Zaharias, J., Finn, J., Folger, J., Johnston, J. and Word, E. (2008). 'Tennessee's student teacher achievement ratio (STAR) project', Harvard Dataverse, V1, https://doi.org/10.7910/DVN/SIWH9F.

Alan, S., Boneva, T. and Ertac, S. (2019). 'Ever failed, try again, succeed better: Results from a randomized educational intervention on grit', *The Quarterly Journal of Economics*, vol. 134(3), pp. 1121–62.

Alan, S. and Ertac, S. (2018). 'Fostering patience in the classroom: Results from randomized educational intervention', *Journal of Political Economy*, vol. 126(5), pp. 1865–911.

Almlund, M., Duckworth, A.L., Heckman, J. and Kautz, T. (2011). 'Personality psychology and economics', in (E.A. Hanushek, S. Machin and L. Woessmann, eds.), *Handbook of the Economics of Education*, vol. 4, pp. 1–81, Amsterdam: Elsevier.

Altonji, J.G., Elder, T.E. and Taber, C.R. (2005). 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of Political Economy*, vol. 113(1), pp. 151–84.

Ballis, B. (2023). 'Dreamers and beyond: Examining the broader educational effects of DACA', *Journal of Human Resources*, doi: 10.3368/jhr.0621-11696R2.

Bertrand, M. and Pan, J. (2013). 'The trouble with boys: Social influences and the gender gap in disruptive behavior', *American Economic Journal: Applied Economics*, vol. 5(1), pp. 32–64.

Bietenbeck, J. (2020). 'The long-term impacts of low-achieving childhood peers: Evidence from Project STAR', *Journal of the European Economic Association*, vol. 18(1), pp. 392–426.

Booij, A.S., Leuven, E. and Oosterbeek, H. (2017). 'Ability peer effects in university: Evidence from a randomized experiment', *The Review of Economic Studies*, vol. 84(2), pp. 547–78.

Borghans, L., Duckworth, A.L., Heckman, J.J. and Ter Weel, B. (2008). 'The economics and psychology of personality traits', *Journal of Human Resources*, vol. 43(4), pp. 972–1059.

Brenoe, A.A. and Zölitz, U. (2019). 'Exposure to more female peers widens the gender gap in STEM participation', *Journal of Labor Economics*, vol. 38(4), pp. 1009–54.

Cadena, B.C. and Keys, B.J. (2015). 'Human capital and the lifetime costs of impatience', *American Economic Journal: Economic Policy*, vol. 7(3), pp. 126–53.

Carrell, S.E. and Hoekstra, M.L. (2010). 'Externalities in the classroom: How children exposed to domestic violence affect everyone's kids', *American Economic Journal: Applied Economics*, vol. 2(1), pp. 211–28.

Carrell, S.E., Hoekstra, M. and Kuka, E. (2018). 'The long-run effects of disruptive peers', *American Economic Review*, vol. 108(11), pp. 3377–415.

Carrell, S.E., Sacerdote, B.I. and West, J.E. (2013). 'From natural variation to optimal policy? The importance of endogenous peer group formation', *Econometrica*, vol. 81(3), pp. 855–82.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W. and Yagan, D. (2011). 'How does your kindergarten classroom affect your earnings? Evidence from Project STAR', *The Quarterly Journal of Economics*, vol. 126(4), pp. 1593–660.

Duckworth, A.L., Peterson, C., Matthews, M.D. and Kelly, D.R. (2007). 'Grit: Perseverance and passion for long-term goals', *Journal of Personality and Social Psychology*, vol. 92(6), pp. 1087–101.

Dunifon, R. and Duncan, G.J. (1998). 'Long-run effects of motivation on labor-market success', *Social Psychology Quarterly*, vol. 61(1), pp. 33–48.

Feld, J. and Zölitz, U. (2017). 'Understanding peer effects: On the nature, estimation and channels of peer effects', *Journal of Labor Economics*, vol. 35(2), pp. 387–428.

Figlio, D.N. (2007). 'Boys named Sue: Disruptive children and their peers', *Education Finance and Policy*, vol. 2(4), pp. 376–94.

Finn, J.D., Boyd-Zaharias, J., Fish, R.M. and Gerber, S.B. (2007). *Project STAR and beyond: Database user's guide*, Tennessee: HEROS, Incorporated.

Gensowski, M. (2018). 'Personality, IQ, and lifetime earnings', *Labour Economics*, vol. 51, pp. 170–83.

Golsteyn, B.H., Grönqvist, H. and Lindahl, L. (2014). 'Adolescent time preferences predict lifetime outcomes', Economic Journal, vol. 124(580), pp. F739–61.

Golsteyn, B.H., Non, A. and Zölitz, U. (2021). 'The impact of peer personality on academic achievement', *Journal of Political Economy*, vol. 129(4), pp. 1052–99.

Guryan, J., Kroft, K. and Notowidigdo, M.J. (2009). 'Peer effects in the workplace: Evidence from random groupings in professional golf tournaments', *American Economic Journal: Applied Economics*, vol. 1(4), pp. 34–68.

Hancock, S.A. and Hill, A.J. (2022). 'The effect of teammate personality on team production', *Labour Economics*, vol. 78, 102248.

Heckman, J.J., Jagelka, T. and Kautz, T.D. (2019). 'Some contributions of economics to the study of personality', Working Paper No. 26459, National Bureau of Economic Research.

Heckman, J., Pinto, R. and Savelyev, P. (2013). 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes', *American Economic Review*, vol. 103(6), pp. 2052–86.

Hoxby, C. (2000). 'Peer effects in the classroom: Learning from gender and race variation', Working Paper No. 7867, National Bureau of Economic Research.

Hoxby, C.M. and Weingarth, G. (2005). 'Taking race out of the equation: School reassignment and the structure of peer effects', Working paper, Harvard University.

Hulleman, C.S. and Barron, K.E. (2015). 'Motivation interventions in education: Bridging theory, research, and practice', in (L. Corno and E.M. Anderman, eds.), *Handbook of Educational Psychology*, pp. 174–85, New York: Routledge.

Kautz, T., Heckman, J.J., Diris, R., Weel, B.T. and Borghans, L. (2014). 'Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success', Working Paper No. 20749, National Bureau of Economic Research.

Krueger, A.B. (1999). 'Experimental estimates of education production functions', *The Quarterly Journal of Economics*, vol. 114(2), pp. 497–532.

Krueger, A.B. and Whitmore, D.M. (2001). 'The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star', Economic Journal, vol. 111(468), pp. 1–28.
Lavy, V., Paserman, M.D. and Schlosser, A. (2012). 'Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom', Economic Journal, vol. 122(559), pp. 208–37.
Lavy, V. and Schlosser, A. (2011). 'Mechanisms and impacts of gender peer effects at school', *American Economic Journal: Applied Economics*, vol. 3(2), pp. 1–33.
Lazowski, R.A. and Hulleman, C.S. (2016). 'Motivation interventions in education: A meta-analytic review', *Review of Educational Research*, vol. 86(2), pp. 602–40.
Mas, A. and Moretti, E. (2009). 'Peers at work', *American Economic Review*, vol. 99(1), pp. 112–45.
Milchus, N.J., Farrah, G.A. and Reitz, W. (1968). *The Self-Concept and Motivation Inventory: What Face Would You Wear?*, Dearborn Heights, MI: Person-O-Metrics.
Murphy, P.K. and Alexander, P.A. (2000). 'A motivated exploration of motivation terminology', *Contemporary Educational Psychology*, vol. 25(1), pp. 3–53.
Oster, E. (2019). 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics*, vol. 37(2), pp. 187–204.
Paloyo, A.R. (2020). 'Peer effects in education: Recent empirical evidence', in (S. Bradley and C. Green, eds.), *The Economics of Education (Second Edition)*, pp. 291–305, London: Academic Press.
Piatek, R. and Pinger, P. (2016). 'Maintaining (locus of) control? Data combination for the identification and inference of factor structure models', *Journal of Applied Econometrics*, vol. 31(4), pp. 734–55.
Poropat, A.E. (2009). 'A meta-analysis of the five-factor model of personality and academic performance', *Psychological Bulletin*, vol. 135(2), pp. 322–38.
Roberts, B.W. (2006). 'Personality development and organizational behavior', *Research in Organizational Behavior*, vol. 27, pp. 1–40.
Roberts, B.W., Harms, P., Smith, J.L., Wood, D. and Webb, M. (2006). 'Using multiple methods in personality psychology', in (M. Eid and E. Diener, eds.), *Handbook of Multimethod Measurement in Psychology*, pp. 321–35, Washington, DC: American Psychological Association.
Roberts, B.W. and Yoon, H.J. (2022). 'Personality psychology', *Annual Review of Psychology*, vol. 73, pp. 489–516.
Romano, J.P. and Wolf, M. (2005a). 'Exact and approximate stepdown methods for multiple hypothesis testing', *Journal of the American Statistical Association*, vol. 100(469), pp. 94–108.
Romano, J.P. and Wolf, M. (2005b). 'Stepwise multiple testing as formalized data snooping', *Econometrica*, vol. 73(4), pp. 1237–82.
Sacerdote, B. (2011). 'Peer effects in education: How might they work, how big are they and how much do we know thus far?', in (E.A. Hanushek, S. Machin and L. Woessmann, eds.), *Handbook of the Economics of Education*, vol. 3, pp. 249–77, Amsterdam: Elsevier.
Schanzenbach, D.W. (2006). 'What have researchers learned from Project STAR?', *Brookings Papers on Education Policy*, vol. 9(9), pp. 205–28.
Segal, C. (2012). 'Working when no one is watching: Motivation, test scores, and economic success', *Management Science*, vol. 58(8), pp. 1438–57.
Shan, X. and Zölitz, U. (2022). 'Peers affect personality development', Discussion Paper No. 15257, IZA (a preprint).
Shavelson, R.J., Hubner, J.J. and Stanton, G.C. (1976). 'Self-concept: Validation of construct interpretations', *Review of Educational Research*, vol. 46(3), pp. 407–41.
Shure, N. (2021). 'Non-cognitive peer effects in secondary education', *Labour Economics*, vol. 73, 102074.
Sojourner, A. (2013). 'Identification of peer effects with missing peer data: Evidence from Project STAR', Economic Journal, vol. 123(569), pp. 574–605.
Sorrenti, G., Zölitz, U., Ribeaud, D. and Eisner, M. (2024). 'The causal impact of socio-emotional skills training on educational success', *The Review of Economic Studies*, doi: 10.1093/restud/rdae018.
Steinmayr, R. and Spinath, B. (2009). 'The importance of motivation as a predictor of school achievement', *Learning and Individual Differences*, vol. 19(1), pp. 80–90.
Whitmore, D. (2005). 'Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment', *American Economic Review*, vol. 95(2), pp. 199–204.
Wong, M.M. and Csikszentmihalyi, M. (1991). 'Motivation and academic achievement: The effects of personality traits and the duality of experience', *Journal of Personality*, vol. 59(3), pp. 539–74.
Word, E., Johnston, J., Bain, H.P., Fulton, D.B., Achilles, C.M., Lintz, M.N., Folger, J. and Breda, C. (1990). 'The state of Tennessee's Student/Teacher Achievement Ratio (STAR) project: Technical report 1985-1990', Project report, Tennessee State Department of Education and others.

The Economic Journal, **135** (*January*), 36–58 https://doi.org/10.1093/ej/ueae060
Advance Access Publication Date: 27 June 2024