The Impact of Peer Personality on Academic Achievement*

Bart H.H. Golsteyn
(Maastricht University)

Arjan Non
(Maastricht University)

Ulf Zölitz
(University of Zurich)

March 2019

Abstract

This paper provides evidence of a novel facet of peer effects by showing that peer personality influences academic achievement. We exploit random assignment of students to university sections and find that students perform better in the presence of persistent peers. This effect is complemented by teacher quality. The impact of peer persistence is enduring, as students exposed to persistent peers at the beginning of their studies continue to achieve higher grades in subsequent periods. The personality peer effects that we document are distinct from other observable peer characteristics and suggest that peer personality traits affect human capital accumulation.

Keywords: peer effects, socio-emotional skills, personality
JEL classification: I21, I24, J24

* We received helpful comments from Armin Falk, Jan Feld, Wolter Hassink, Ingo Isphording, Fabian Kosse, Nico Rasters, Nicolas Salamanca, Giuseppe Sorrenti, Tom Stolp, Dirk Tempelaar, seminar participants at briq, the University of Cologne, the University of Munich, Maastricht University, Utrecht University, the Max Planck Institute for Research on Collective Goods (Bonn), the University of Bamberg, the University of Zurich, HEC Lausanne, the University of Reading, the University of Sheffield, and conference participants at the briq/IZA workshop on socio-emotional skills, the 2nd IZA education workshop, the workshop on Measurement of Individual Heterogeneity and its Use in Economic Models (Bonn), the European Economic Association (Cologne), the European Society for Population Economics (Antwerp), the International Workshop on Applied Economics of Education (Catanzaro), the Institute for the Study of Labor/Society of Labor Economists (Buch am Ammersee), the Economics of Education workshop (Maastricht), the Royal Economic Society (Sussex), the European Association of Labour Economists (Lyon), and the Society of Labor Economists (Toronto). Golsteyn acknowledges financial support from the Netherlands Organization for Scientific Research (NWO VIDI grant 452-16-006). We thank Sophia Wagner and Jeffrey Yusof for providing outstanding research assistance and the Department of Education Research and Development for sharing the data. Bart H.H. Golsteyn: Maastricht University, Department of Economics. b.golsteyn@maastrichtuniversity.nl; Arjan Non: Maastricht University, Department of Economics. j.non@maastrichtuniversity.nl; Ulf Zölitz: University of Zurich, Department of Economics & Jacobs Center for Productive Youth Development, IZA, CESifo, CEPR and Maastricht University. ulf.zoelitz@econ.uzh.ch.

I. Introduction

A steadily growing body of literature in economics, psychology, and sociology recognizes the importance of people's personality traits on their life's trajectory. Personality traits predict many significant outcomes in life, including educational attainment, earnings, employment, and health as well as participation in risky behavior and crime.[1] Although evidence on the importance of personality traits is accumulating, there is virtually no evidence on the extent to which individuals' personality traits affect other people in their social environment. The literature on peer effects, which is dedicated to identifying social spillovers, has established that peer characteristics such as race, gender, and test scores affect the accumulation of human capital.[2] Surprisingly, the literature neglects the question of whether the *personality* of peers affects educational outcomes.

In this paper, we test whether peer personality affects performance in education using data from a Dutch business school. Two key features make this institution the ideal place to study peer effects. First, students are required to spend a significant amount of study time in small teaching sections of up to sixteen students. These sections provide us with natural peer groups in which students engage in meaningful social interactions through solving problems and discussing the literature. Second, students are randomly assigned to these sections, conditional on scheduling constraints. In most other educational settings, students are tracked or self-select into peer groups. Random assignment allows us to overcome this fundamental selection problem that typically plagues the identification of causal peer effects.

---

[1] An often-used taxonomy in personality psychology is the Big Five. This entails openness to experiences, conscientiousness, extraversion, agreeableness, and neuroticism. Big Five traits are predictive of many outcomes in life, including schooling, wages, crime, teenage pregnancy, and longevity. Generally, conscientiousness and neuroticism are more predictive than the other traits (see Borghans et al. 2008, Almlund et al. 2011). Borghans et al. (2016) show that personality traits have predictive power over and above pure cognition for life outcomes. Regarding economic preference parameters, recent studies reveal that time preferences (Golsteyn, Grönqvist and Lindahl 2014, Åkerlund et al. 2016), risk attitudes (Dohmen et al. 2011), and social preferences (Dohmen et al. 2009) predict outcomes in life such as educational attainment, wages, and health outcomes.

[2] Sacerdote (2011) provides a review of the literature.

We measure personality at the beginning of the students' academic careers, before their assignment to the groups in which we test for peer effects. This avoids the simultaneity problem arising from the fact that peers may influence students' self-reported personality. We collect four distinct measures of students' personality traits related to education: (1) *persistence*, which we show is a facet of grit (Duckworth et al. 2007) and which captures how much students keep trying to solve a problem even if it is challenging; (2) *self-confidence*, which measures students' belief in their ability to do well and succeed in their studies; (3) *anxiety*, which we show is a facet of neuroticism (a "Big Five" personality factor) and reflects how worried and nervous students are about succeeding in their studies; and (4) *risk attitude*, which captures people's willingness to engage in risky behaviors. After combining students' self-assessed measures of these traits with their administrative records on grades, we observe 17,512 student-course observations of 4,383 students from five study cohorts.

Our results show that students who were randomly assigned to a group of more-persistent peers attain higher exam grades in centrally graded exams. A one standard deviation increase in average peer persistence raises grades by 1.8 percent of a standard deviation. The impact of a one standard deviation increase in peer persistence is approximately twice as large as a one standard deviation increase of peer GPA as identified by Feld and Zölitz (2017) in the same setting. We further find that exposure to risk-tolerant peers negatively affects performance. A one standard deviation increase in peers' risk tolerance lowers grades by 1.1 percent of a standard deviation. Peer anxiety and self-confidence do not significantly affect performance. While the impact of peer persistence remains statistically significant after correcting for multiple hypothesis testing, the impact of peer risk attitude on performance does not remain significant after the correction.

Given that personality traits are likely to be correlated with other characteristics, we test whether the inclusion of peer GPA, peer math and statistics entry test scores, peer gender, and peer nationality affects our results. Our point estimates on the impact of peer personality are not affected by the inclusion of these observable peer characteristics. These results suggest that the effect of peer personality is distinct from peer achievement and other observable peer characteristics.

What are the mechanisms behind these effects? Persistent peers may raise performance, for example, through influencing study efforts, through improving group functioning, or through motivating teachers. To better understand the underlying mechanisms, we analyze: (1) students' self-reported study hours; (2) students' evaluations of courses and teachers; (3) how peer personality interacts with teacher quality; and (4) whether performance in parallel and subsequent courses is affected. While we do not find that persistent peers affect self-reported study hours or students' perceptions of the teacher, we do find that peer personality is complementary to teacher quality. The impact of peer persistence on performance is particularly pronounced when the group is taught by a high value-added teacher. Moreover, we find that the effect of persistent peers is nontransient: exposure to persistent peers also increases grades in future periods. This suggests that our results are not driven by a transfer of course-specific knowledge, but rather by an accumulation of human or social capital: students with access to more-persistent peers may acquire more skills or build a friendship network from which they benefit in current and future courses.

When looking at the heterogeneity of peer effects, we find that both students with high and low own-persistence benefit from having persistent peers. Students with a middle level of persistence are not affected. These effects suggest that grouping high- and low-persistence students together could be an effective way to reduce inequality in academic achievement.[3]

[3] Note that Carrell, Sacerdote, and West (2013) caution against this type of policy recommendation, because: (1) we may not have sufficient underlying support in the data, and (2) the reassignment may change the underlying structure

The evidence on externalities arising from peer personality on educational outcomes is scarce. Most of the literature on peer effects in education focuses on how various measures of peer achievement affect performance. Papers that exploit random assignment to identify a causal effect of peer achievement include Zimmerman (2003), Whitmore (2005), Carrell, Fullerton, and West (2009), Lyle (2009), Duflo, Dupas, and Kremer (2011), Carrell, Sacerdote, and West (2013), De Giorgi and Pellizzari (2013), and Booij, Leuven, and Oosterbeek (2017). Feld and Zölitz (2017) study achievement peer effects in the same setting as our present paper.[4] Prominent papers studying the effects of peer race and gender include Hoxby (2000), Angrist and Lang (2004), Hoxby and Weingarth (2005), Lavy and Schlosser (2011), Oosterbeek and van Ewijk (2014), and Einiö and d'Este (2018).

Only a few papers have touched upon the effect of peer characteristics related to personality. Figlio (2007) shows that boys with female-sounding names, who are generally more likely to display behavioral problems, negatively affect their peers' test scores. Carrell and Hoekstra (2010) show that children who experienced domestic violence negatively impact their peers' reading and math test scores. Carrell, Hoekstra, and Kuka (2018) follow up on these results and show that exposure to disruptive peers in elementary school reduces earnings at the age of 26 by 3 to 4 percent. While the role of peer personality is not explicitly studied in any of these papers, their findings suggest that personality traits that underlie students' disruptive behavior negatively affect their classroom peers.

of social interactions between students that generate peer effects. An experimental validation of the policy implications that we suggest here would provide important additional insights into whether our results can be generalized.

[4] Note that although Feld and Zölitz (2017) use data from the same institution, their estimation sample does not overlap with ours. Feld and Zölitz (2017) use administrative data from 2009–2011. In contrast, we combine administrative data from 2012–2016 with personality data we collected using surveys among all incoming students in these five years.

The only paper to explicitly consider the association between peer personality and educational outcomes is by Shure (2017), who investigates the relationship between peer personality and grades in secondary school. In order to identify peer effects, she exploits idiosyncratic variation in peer personality in a school fixed effects framework. Consistent with our results, Shure (2017) documents a positive relationship between peer conscientiousness and student performance. Moreover, she finds a negative association between performance and peer extraversion. The key difference between Shure (2017) and our study lies in the empirical strategy: while Shure (2017) relies on school fixed effects, we exploit the random assignment of students into teaching sections. Our approach alleviates concerns related to nonrandom student sorting into peer groups.

Our paper establishes a novel connection between the peer effects literature and a growing body of evidence documenting the importance of individuals' personality traits. We make two contributions to these two strands of literature. First, by studying the role of peer personality instead of achievement, gender, or race, we focus on a novel facet of peer effects in human interaction. Second, we provide the first causal evidence that personality traits of people around us affect educational performance.

Our key finding that peer personality affects performance has important implications for the social returns of interventions that affect socio-emotional skills. For instance, the social returns of interventions like the Perry Preschool Project, which enhanced socio-emotional skills, will be underestimated if spillovers driven by peer personality are neglected. Such spillovers may arise, for instance, if children affected by the intervention in turn affect their siblings, friends, neighbors, or classmates.

The remainder of the paper is structured as follows. Section II describes the institutional environment and the assignment procedure of students to sections. Section III describes the data set. Section IV discusses the empirical strategy and shows evidence that section assignment is random. Section V reports results. Section VI investigates underlying channels. Section VII concludes.

II. Institutional Background

We collected the data from a business school in the Netherlands.[5] At present, about 4,300 students are enrolled in bachelor's, master's, and PhD programs at this institution. In contrast to the US college system, all students who enroll at the business school are committed to studying a specific program from the first year onward. In all bachelor's programs, students must take eight compulsory courses and two shorter skills courses in the first year. Some of those courses are program-specific. In the second and third years, students choose a major and elective courses. In this paper, we concentrate on performance in the first year, when all students in a given study program take a set of compulsory courses and take centrally organized and mostly machine-graded exams.

The academic year at the business school is divided into four regular teaching periods of two months each and two two-week skills periods. Students usually take two courses simultaneously in each regular period and one course in each skills period. In our analysis, we

[5] For similar but more detailed information on the institutional environment, see Feld, Salamanca, and Hamermesh (2016), Feld and Zölitz (2017), and Feld, Salamanca, and Zölitz (forthcoming).

focus on the courses taken during the regular teaching periods because students are often not graded in skills courses or the relevant peer group cannot be identified.[6]

Importantly, the bulk of teaching at the business school occurs in sections. Sections are small groups of up to sixteen students, and are taught by one instructor. This peer group will be the focus of our analysis. Students typically meet twice per week for two hours with their section peer group to discuss the course material. In this discussion-based approach, students generate questions about a topic, try to answer these questions in self-study, and then discuss their findings with their peers in the next session. The role of the instructor in this discussion-based system is to monitor and guide the classroom discussion. In most courses, students solve problem sets and read course materials at home and then meet to discuss the material and solutions to the problem sets. For each course, attendance of the majority of section meetings is compulsory. Additionally, most courses have lectures that students of all sections attend.

To assess the external validity of our findings, it is important to know how prevalent small-group teaching is. Feld, Salamanca, and Zölitz (forthcoming) shed light on this question by collecting information on the prevalence of small-group teaching using a survey of 69 economics and business university departments in 31 OECD countries. The results of this survey show that the majority—63 percent—of OECD universities use small-group teaching. At 66 percent of those universities, students discuss course material and exercise solutions in small group meetings. Small group meetings make up around 30 percent of students' contact hours, and the average group size is 16.3 students. In our sample, the average group size is similar, containing 12.9 students per section on average. These results show that many OECD universities also use small-group teaching, which supports the external validity of our results.

[6] In almost all skills courses, students are scheduled in different sections but end up sitting together in the same room. Furthermore, some skills courses have only a "pass" or a "fail" grade.

A. Assignment of Students to Sections

The Scheduling Department of the business school assigns students to sections, teachers to sections, and allocates sections to time slots and rooms. The department uses a computer program that randomly allocates all registered students for a given course into sections. Undergraduate students are further stratified by nationality.[7] After students are assigned to sections, teachers are assigned to sections, and then sections are assigned to available time slots and rooms.[8] After this assignment, the scheduling program indicates scheduling conflicts.[9] These conflicts arise for about 5 percent of the initial assignments. If the computer program indicates a scheduling conflict, the scheduler manually moves students between different sections until all scheduling conflicts are resolved. After this, the section and teacher assignments are published.

Schedulers have little interaction with students and therefore do not know them personally. They schedule thousands of students every eight weeks and do not have time to pay special attention to individual students. The scheduling program does not show students' previous grades, gender, or personality traits. Schedulers follow the random allocation mechanism that the

[7] This was introduced in the academic year 2010/2011. The stratification occurs as follows. The scheduler first selects all German students (who are not ordered by any observable characteristic) and then uses the option "Allocate Students set SPREAD," which assigns an equal number of German students to all sections. Subsequently, the scheduler repeats this process with the Dutch students, and finally distributes the students of all other nationalities into the remaining spots. Until the academic year 2013/14, about 10 percent of the slots in each section were initially left empty and were filled with students who registered late. This procedure balanced the number of late registration students over the sections. Since 2013/14, the institution no longer admits students to courses after the registration deadline. Starting with the academic year 2013/14, students are automatically registered for first-year bachelor's courses.

[8] About 10 percent of teachers indicate time slots when they are not available for teaching. This happens before they are scheduled and requires the signature of the department chair.

[9] There are three reasons for students' scheduling conflicts: (1) the student takes another elective course at the same time; (2) the student is also working as a teaching assistant and is scheduled to teach at the same time; or (3) the student indicates nonavailability for evening education. By default, all students are recorded as available for evening sessions, but they can opt out by completing an online form. About three percent of all sessions in our sample are scheduled for an evening time slot.

scheduling software suggests and only ensure that the two main nationalities (Dutch and German) are somewhat evenly distributed over teaching sections. Students do not interfere with the scheduling process. Formal switching between assigned sections is not possible, and section instructors are instructed not to allow informal switching.

There are a few exceptions to this general procedure, e.g. when the course coordinator requests to manipulate the section composition. This is the case for only one of the first-year courses, which we therefore remove from the estimation sample. Importantly, in the estimation sample that we use throughout this paper, neither teachers, students, nor course coordinators influence the section assignment. We formally test whether the data have the characteristics expected under random assignment in section IV.

III. Data

A. Sample and Descriptive Statistics

For the academic years 2012/13 through 2016/17, we collected data on students' personality using online questionnaires that students were required to complete at the beginning of their introductory course in quantitative methods. This course takes place in the first period of the first academic year and is obligatory for all economics and business students at the business school. Because the survey was part of a compulsory assignment students completed for the course, virtually all students filled out the questionnaires.[10] Only a handful of students who dropped out of the study program during the first weeks did not answer the survey. Our sample thus comprises five full study cohorts.[11]

Panel A in Table 1 provides a data overview. In total, we observe 4,383 first-year bachelor's students.[12] For the 17,512 student-course registrations, we observe 16,155 student-course grades. On average, 8 percent drop out of a course between registration and sitting the final exam. Students who dropped out of courses are included in the analysis as peers. Therefore, the peer effects we estimate represent intention to treat effects. The total number of sections is 1,357. There are 23.4 sections per course on average, and the average number of students per section is 12.90.

Panel B in Table 1 shows student demographic characteristics and personality measures. The panel shows that around 40 percent of the students are female. As the business school is located close to the German border, almost 50 percent of the students are German, while 24 percent of students are Dutch. Panel B also provides summary statistics for our measures of student personality that we will describe in section C below. Panel C in Table 1 summarizes the section-level peer characteristics that we construct to estimate peer effects.

---

[10] Students were informed that their responses would remain confidential and that they would be used for research purposes as well as the general improvement of education.

[11] The Scheduling Department of the business school provided the administrative data on all scheduled sections. The Examinations Office of the business school provided the data on student course registrations, grades, and student background characteristics.

[12] The academic year consists of four periods, with two courses taught each period. In our analyses, we exclude the two courses in the first course period to avoid the reflection problem (Manski 1993), as well as one course in which the instructor interfered with the randomization procedure, leaving us with five courses. The sample contains 17,512 student-course registrations. This implies that students take four courses on average, which is less than five courses, because some students quit their studies during the first year and therefore are not registered for all five courses. The number of observations in the main analyses equals the number of student-course grades.

B. Data on Student Performance and Student Course Evaluations

The performance indicator in this study is the grade that students achieve in the centrally graded exam at the end of each course. Each student takes the exam, and it does not have a group-graded component. We only use the results of the first central exam in a course and do not take the grades of any exam re-takes into consideration, as these are not comparable to the grades in the first take. Panel D in Table 1 shows that the average student grade is 6.5.[13]

In our analyses of mechanisms, we use survey data from online course evaluations. Panel D shows the evaluation items we analyze. Students report that they study on average around thirteen hours per week for one course. This excludes the six hours per week during which they attend sections and lectures. Given that the students take two courses per period, this number of study hours is close to the 40 hours per week that full-time students are supposed to invest in their studies according to the European Credit Transfer System (ECTS) framework.

The evaluation questions regarding the instructor were: "The tutor sufficiently mastered the course content," "The tutor stimulated the transfer of what I learned in this course to other contexts," "The tutor was enthusiastic in guiding our group," and "Evaluate the overall functioning of your tutor in this course with a grade." The evaluation questions about peer-to-peer interaction were: "Working in tutorial groups with my fellow students helped me to better understand the subject matter," and "My tutorial group has functioned well." For the purpose of our analysis, we aggregate the answers to the four questions about the instructor and the answers to the two questions about peers into two indices. To construct these indices, we standardize the separate items, compute the average of peer- and instructor-related items respectively, and standardize again.

[13] The Dutch grading scale ranges from 1 to 10, with 5.5 usually being the lowest passing grade. If the grade of a student is lower than 5.5, the student fails the course and has the opportunity to take the exam a second and third time.

C. Measures of Student Personality and Attitudes

Table 2 provides an overview of the personality measures that we use.[14] All measures are self-reported on a scale from 1 to 7. We use the Student Motivation Scale as introduced by Martin (2009) to measure persistence, self-confidence, and anxiety. Each of the traits is measured by four questions. This scale is specifically developed for the measurement of student motivation in education. Hence, all questions are framed in the context of education.

How reliable are the measures we use and how much do these personality measures overlap with other, more commonly used measures of personality? In Appendix B we show that the test-retest correlation of these traits is fairly high, ranging between 0.57 and 0.75. Appendix B also provides evidence that our measure of persistence is strongly correlated with conscientiousness, and that it is a facet of grit (Duckworth et al. 2007). Furthermore, Appendix B shows that our measure of anxiety is a facet of neuroticism.

Our measure of risk attitudes is the widely used question: "In general, how willing are you to take risks?" Higher values indicate higher risk tolerance. Dohmen et al. (2011) and Vieider et al. (2015) show that this measure predicts behavior in incentivized lottery experiments, and that it is correlated with risky behaviors in several domains across different cultures (see also Falk et al. 2016).

[14] Table A1 in the appendix shows the correlations between these variables.

IV. Empirical Strategy and Randomization Check

A. Empirical Strategy

Our goal in this paper is to estimate the effect of peer personality on students' performance. Throughout the paper, we define peer groups at the section level; when referring to peers, we mean students' section peers. Before we test how peer personality affects outcomes, we investigate whether students' *own* personality traits predict their outcomes. We estimate the following model:

$$GPA_i = P_i \, \delta' + X_i \theta' + \varepsilon_i, \qquad (4)$$

where GPA_i is the grade point average of student *i* at the end of the first study year. The vector of personality traits P_i includes the student's persistence, self-confidence, anxiety, and risk attitude. The effect of these traits on student performance is captured by the vector δ'. X_i is a vector of control variables that include student gender, nationality, and year-times-study-program fixed effects.

To test how peer personality affects student performance, we estimate the following model:

$$A_{icg} = \overline{PP}_{g-i} \, \alpha' + P_i \, \beta' + X_i \gamma' + \rho_{ic} + u_{ic}, \qquad (5)$$

where A_{icg} is the grade of student i in course c in section g. The vector $\overline{PP}_{g-i}$ refers to the mean personality traits of all students in section g excluding the individual itself, i.e. the leave-out mean. We control for several variables to enhance the precision of our estimates: students' own personality measures and a vector of other control variables X_i that include students' own gender and nationality, their GPA at the start of the course, and indicators for scheduling conflicts. We include the latter to account for potential nonrandom assignment due to scheduling conflicts. ρ_{ic} denotes course-year fixed effects and u_{ic} is the error term. We follow Abadie et al. (2017) and cluster standard errors at the level of randomization, which is at the course level in our case.[15]

Personality traits are correlated with GPA. Therefore, one obvious question is whether α' in equation (5) is indeed picking up the true effect of peer personality on student performance or whether α' is picking up an underlying achievement peer effect. To assess to what degree this is the case, we also estimate an extended version of this model in which we control for other peer characteristics such as GPA, gender, and nationality to disentangle the impact of peer personality from these characteristics.[16] In a further robustness check, we include peers' entry test scores for math and statistics. The difference in the peer personality coefficients in the model with and without these controls provides information on the extent to which these observables affect the relationship between peer personality and performance. If the estimates remain robust to the inclusion of these observables, it is likely that they also remain robust to factors not included in the estimations (see Altonji, Elder, and Taber 2005). To investigate heterogeneous treatment effects, we also estimate a variant of (5) where we allow α' to vary by students' own level of the respective trait.

Feld and Zölitz (2017) have shown that classical measurement error in the peer characteristics of interest can lead to substantial overestimation of peer effects when group assignment is nonrandom. When group assignment is random—as is the case in our setting—classical measurement error will attenuate peer effects estimates, i.e. bias them toward zero.

[15] The results do not change when we cluster on the section level (see Table A6 in the Appendix).

[16] Peer GPA is defined as the running average of grades obtained in all *previous* courses. This implies that whenever we include GPA in an estimation, we refer to a pre-treatment characteristic that was fixed prior to the randomization into sections that takes places before the start of each new course. Peer GPA can therefore not be affected by students' own characteristics.

Because peer personality is arguably measured with a substantial amount of error, we expect that our estimates of α' will be significantly attenuated. This implies that we identify lower bounds and that the true underlying effect is likely to be larger than our estimates.

In order to simplify the interpretation of our estimates, we standardize own and peer personality measures as well as course grades to have a mean of zero and unit variance.

B. Tests for Random Assignment

The key identifying assumption of this paper is that the assignment of students to sections (i.e. peer groups) is random. The scheduling procedure described in Section III ensures that student assignment to sections is random, conditional on scheduling conflicts. Using data from the same environment, Feld and Zölitz (2017) have shown that section assignment has the properties that one would expect under random assignment. To confirm this result with respect to the sample we study in this paper and with respect to peer personality, we test whether student personality relates to average peer personality in the assigned section. This randomization check closely follows Guryan, Kroft, and Notowidigdo (2009) and controls for the course-level leave-out mean of the respective characteristic to account for the mechanical relationship between own- and peer-level variables. Table 3 reports the results of this analysis and shows that peer personality is not systematically related to students' own personality. All coefficients are small and not statistically significant, which confirms that the section assignment is random.[17]

[17] The number of observations in the randomization check differs from the number of observations in the main analyses because some students drop out of the courses. Table A3 in the appendix provides balancing tests for the sample of students who did not drop out and shows that students' pretreatment characteristics are also unrelated to peer personality when conditioning on this subsample. Table A2 shows that peer personality does not affect the probability of dropping out of a course.

C. Simultaneity Problem

A potential threat to identification of personality peer effects occurs when personality and academic performance are measured simultaneously. Simultaneous measurement may induce a correlation between peer personality and performance, as students could influence each other's assessments and face the same shocks. This may lead to bias in the estimation of peer effects. In our data, this could occur in the first course period at university, but not in later periods because students are randomly reassigned to a new section of peers in each of these periods. By excluding the first course period from our estimation sample, we ensure that our estimates are not affected by the simultaneity problem.

V. Results

A. The Relationship between Students' Own Personality and Performance

Are personality traits relevant predictors of students' own performance? We investigate this question by looking at how students' own personality traits, as measured at the beginning of the first study year, relate to student GPA at the end of their first year in university. Table 4 shows that all personality traits that we measure are significantly correlated with GPA. A one standard deviation increase in persistence is correlated with 0.13 standard deviations higher GPA. Self-confidence is also positively related to GPA with a similar magnitude. We further find that anxiety is negatively related to GPA. A one standard deviation increase in anxiety is associated with a 0.16 standard deviation reduction in GPA. We also find that students who are more risk tolerant have a lower GPA. A one standard deviation increase in risk tolerance is related to a 0.09 standard deviations decrease in GPA. For reference purposes, we also estimate a regression of GPA on a dummy variable that equals one if a student has a high school math major. As seen in column (5), the size of the personality coefficients is roughly equal to half of the size of the high school math major indicator. In column (6), we include all personality measures in one model. While the magnitudes of the estimated coefficients change by some degree, they remain highly statistically significant and do not change sign.

Figure 1 provides an illustration of these relationships. The plots in Figure 1 visualize the regression results reported in column (6) of Table 4. The construction of these binned scatter plots follows Chetty, Friedman, and Rockoff (2014). We first regress course grades on the set of controls included in column (6) of Table 4 to obtain the residualized course grades. Next, we rank-order observations by our measures of personality and split them into equally sized bins. Subsequently, we plot the mean of the residualized course grades within each bin against the normalized mean value of personality in that bin. Figure 1 shows that the relationships found in Table 4 are linear.

Taken together, the results in Table 4 and Figure 1 show that students' own personality traits are relevant predictors of study success. Our findings are broadly consistent with previous work on the relationship between educational attainment and personality (see Borghans et al. 2008 and Borghans et al. 2016 for reviews).[18]

B. The Impact of Peer Personality on Performance

Prior to analyzing the effect of peer personality on grades, it is important to test whether peer personality affects first-year course dropout. The estimation results, reported in Appendix Table

[18] Borghans et al. (2008) show in their overview of the literature that conscientiousness, which is related to our measure of persistence, is by far the best predictor of grades among the personality traits (r = .22), and that, after openness to experience, it is the best predictor of years of education (r = .11).

A2, show that this is not the case. Column (7) shows that the probability of dropping out of the courses, which determines whether we observe students' grades, is independent of peer personality.

Table 5 displays the estimation results of our main analysis on how peer personality affects students' grades. We find that students who were randomly assigned to more-persistent peers obtain higher exam grades. A one standard deviation increase in peer persistence raises performance by 1.9 percent of a standard deviation. We also find that exposure to risk-tolerant peers negatively affects grades. A one standard deviation increase in peers' risk tolerance lowers grades by 1.1 percent of a standard deviation. We do not find that peers' self-confidence or anxiety are significantly related to grades.

As our measures of peer personality might be collinear to some degree, we include all peer personality variables in one model in column (5) of Table 5. Importantly, the point estimates remain very similar when we include all peer personality measures at once instead of estimating models with one peer characteristic at a time. This suggests that our measures of peer personality capture distinct components of students' personality traits. In column (6), we additionally include fixed effects for scheduling conflicts. The estimated coefficients remain virtually unchanged when we include these fixed effects.

In column (7) of Table 5, we additionally include other observable peer characteristics as control variables. In particular, we include peer GPA, the proportion of female peers, the proportion of Dutch and German peers, as well as the percentage of peers who have a high school math major. If peer personality affects performance mainly through these peer characteristics, we would expect that their inclusion reduces the effect size of the peer personality coefficients. Column (7) shows that this is not the case and that the point estimates remain almost unchanged when we control for other peer observables. In the spirit of Altonji, Elder, and Taber (2005), this result suggests that omitted variables bias plays a limited role in our estimations.

To investigate whether measurement error in peer achievement (GPA) may confound the estimates of peer personality, we control for multiple other measures of peer achievement assessed when students enter the business school, including whether they were a math major in high school, and their math and statistics entry test scores. Table A4 shows that the results remain similar when we include these measures next to peer GPA in the regression. We conclude that it is unlikely that measurement error in peer GPA is confounding the effect of peer personality on grades found in this paper.

In Table 5, we estimate peer effects for four different personality traits. This implies that some statistically significant effects might simply represent chance findings. We address this concern by: (1) performing an F-test for joint significance of all peer personality characteristics, and (2) testing which estimates remain significant after correcting for multiple hypothesis testing. Table 5 shows that we can clearly reject that performance is independent from peer personality. The p-value for the test of joint significance in our most conservative specification in column (7) is .0043. Table A5 further shows that the effect of peer persistence remains statistically significant when using the Bonferroni correction or other commonly used correction methods for multiple hypothesis testing. The effect of peer risk attitude, however, is no longer statistically significant when correcting for multiple hypothesis testing.

In our analyses, we have 58 clusters, i.e. course-times-year observations. A potential threat to statistical inference is that this number of clusters is relatively small. To address this concern, we use the method developed by Donald and Lang (2007), which is a two-step estimator that adjusts standard errors for a small number of clusters. By doing so, we essentially recognize that the fundamental unit of observation is a cluster and not an individual within a cluster. The results reveal that none of our findings change in a qualitative sense when correcting p-values for the small number of clusters. Likewise, the results do not change when we cluster on the section level (see Appendix Table A6).

Taken together, the point estimates presented in Table 5 show that peer persistence has a robust causal impact on student achievement and that personality peer effects are distinct from the effects of peer achievement, gender, or country of origin. Figure 2 visualizes these findings and shows that the effects of peer persistence and peer risk tolerance on performance are fairly linear over the range of available support.

How does the size of the peer persistence effect we find compare to the effects of other peer characteristics and the impact of instructors in the same setting? When we compare our effect to achievement peer effects, we find that the impact of a one standard deviation increase in peer persistence is about twice as large as the impact of a one standard deviation increase in peer GPA (Feld and Zölitz 2017). When compared to gender peer effects, the effect of a one standard deviation increase in peer persistence is about twice as large as the impact of a 10 percentage point increase in the proportion of female peers for women's grades (Zölitz and Feld 2019). Compared to rank effects, the effect of one standard deviation higher peer persistence is about as large as the effect of increasing students' GPA rank by 10 percentiles, i.e. one position in a group of 11 students (Elsner, Isphording, and Zölitz 2018). Comparing our effect to teacher effects shows that the impact of a one standard deviation increase in peers' persistence is approximately the same size as that of being taught by a one standard deviation better teacher as captured by teacher value-added (Feld, Salamanca, Zölitz, forthcoming).

C. Effects on Other Aspects of Performance

Does peer persistence affect other performance aspects besides contemporaneous grades? Table 6 shows how peer personality affects other aggregate study outcomes in the first year. In this analysis, we investigate the overall number of courses passed, successful completion of first-year requirements, and the final first year GPA across all courses. Because these outcomes are invariant at the individual level, we aggregate peer personality across all peers that students meet in first-year sections. This approach reduces the underlying support and statistical power but provides us with an idea of the overall importance of all peers in the first year. Because this approach prevents us from controlling for the exact level of randomization by including course-times-year fixed effects, we instead include year-times-study-program fixed effects to account for differences in the pool of peers. Table 6 shows that peers across all sections of the first year have a meaningful impact on the number of courses passed, on the probability of passing the first year, and the GPA at the end of the first year. Column (1) shows that exposure to more-persistent and self-confident peers raises the course passing rate. Column (2) shows that students with more-persistent peers are also more likely to meet the study requirements for successfully completing the first year.[19] A one standard deviation increase in peer persistence raises the probability of successfully completing the first year by 1.3 percentage points. The effect, however, is only marginally significant. Column (3) shows that having more-persistent peers also raises the GPA of all courses in the first year. A one standard deviation increase in persistence of all peers raises the overall GPA by 2.6 percent of a standard deviation.

[19] Students enrolled in any study program are required to have obtained at least 47 credits (34 credits before 2015) within the first-year bachelor's courses to be eligible for admission to the final bachelor's exam program. At least 6.5 credits of these 47 credits must be obtained within the "Quantitative Methods 1" or "Quantitative Methods 2" courses. Exemptions from these requirements can be granted for medical reasons. First-year completion requirements differ across years and study programs, but typically require passing one or both statistics courses and between two and four additional courses.

VI. Mechanisms

How does the personality of peers affect students' grades? In this section, we analyze the effects of peer personality on students' self-reported study hours and students' evaluations of the course and the teacher. We also investigate the interaction between peer personality and measures of teacher quality and whether performance in parallel and subsequent courses is affected.

A. Self-Reported Study Hours

Peer personality might affect student grades if students start working harder and studying more hours when they work with persistent peers. Using individual-level data from students' course evaluations, we can shed light on this potential explanation. At the end of each period, before learning their grade, students complete an online course evaluation. In these evaluations, students report their average weekly course study hours and evaluate their peer-to-peer interactions, the instructor, and overall course quality.[20]

Table 7 shows estimates for the effects of peer personality on study hours.[21] Column (1) shows that peer persistence does not significantly affect self-reported study hours. The effect is small and precisely estimated. This suggests that increased study hours do not drive the effect of peer persistence on grades. Because study hours may not accurately capture study effort, it remains

[20] Because participation in the questionnaire is voluntary, not all students complete the course evaluation forms. In Appendix Table A7, column (2), we find no evidence that peer personality affects students' probability of responding to the student course evaluation survey. Appendix Table A8 replicates our main analysis for the subsample of students taking part in the course evaluations and shows that point estimates for this subsample and the full sample are very similar.

[21] Our measure of study hours is self-reported, which implies that it may not accurately measure study effort. In order to investigate what exactly study hours measure, we analyze the relationship between own personality traits and own study hours. Table A9 in the Appendix shows that own personality traits predict own study hours. For example, a one standard deviation increase in persistence is related to 1.5 hours more self-reported study time. We find positive effects for anxiety, while self-confidence and risk tolerance are negatively related to study hours. Taken together, Table A9 shows that our measure of study hours is systematically correlated with personality.

possible, however, that exposure to persistent peers raises students' study effectiveness. Our analysis of average study hours therefore does not rule out that exposure to more-persistent peers makes students become more-efficient learners.

B. Teachers

Peer personality might affect student grades if more-persistent peers motivate teachers to provide better or more-tailored instructions. Table 7 provides an indirect test for this mechanism: we test whether peer persistence affects how students perceived the teacher in their teaching evaluations. Column (2) shows that peer persistence does not affect how students evaluate their instructors.

Another mechanism for how peer personality affects grades could be that students with certain personality traits interact more with each other or make the course more attractive for other students. The results in columns (3) and (4) show, however, that peer persistence is also not related to perceptions of group functioning or the overall course quality.[22]

Ideally, we would like to measure teacher quality by value-added measures. We can compute these measures based on students' grades. However, we cannot directly test whether peer personality affects teachers' performance based on such value-added measures. If we used residualized grades to estimate teacher productivity with teacher value-added models, we would falsely attribute the contribution of persistent peers to the teacher and therefore falsely conclude that teachers who teach persistent students have a high value-added. Directly estimating whether

---

[22] Interestingly, we do find some significant effects for peer anxiety in this table. In our main analyses, we find that having anxious peers does not affect grades unless they are *very* anxious. Here, we find that peer anxiety is negatively related to perceptions of group functioning, teacher quality, or the overall course quality. An explanation for these findings may be that anxious peers destroy the classroom atmosphere. Very anxious students may spread their fear of failure to other students, who start questioning the instructor and course quality as a whole.

persistent peers leads to higher teacher value-added would thus provide biased estimates and is therefore not a feasible solution.

What we *can* do is analyze whether the effects of peer personality are moderated by teacher productivity. We test whether peer personality interacts with teacher value-added measures that are constructed based on teacher effectiveness in other courses.[23] If teacher and peer quality are substitutes, we would expect peer personality to be more important in classes with a low-quality teacher. It would suggest that peer-to-peer instruction is an important mechanism in our setting. When teachers are unable to explain the material, persistent peers might be able to compensate for a lack of teacher quality. If, however, teachers and peers are complements, we would expect peer personality to matter more in classes with high-quality teachers. In this case, a good teacher might be able to foster the effect that persistent peers have and in this way enhance student performance. Surprisingly, interaction effects between teacher and peer characteristics have not received any attention in the existing peer effects literature.

In Table 8, we estimate the relationship between peer personality and grades for three subgroups: sections with high, medium, and low teacher quality. Column (1) in Table 8 shows that students with high-quality teachers benefit from having persistent peers, while columns (2) and (3) show that this relationship is not significant for students with lower quality teachers. Taken together, Table 8 provides suggestive evidence on the complementarity of teacher quality and peer persistence. Students benefit from having persistent peers only when they also have good teachers. We can only speculate why this is the case. Perhaps high-quality teachers are more able to adjust their instruction or practices to the classroom composition, and by doing so, they bring out the best in groups with persistent students.

[23] For the construction of teacher value-added models, we leave out the course in which we test for personality peer effects to avoid the possibility that measures of teacher productivity pick up performance gain from peer personality.

C. Effects on Parallel and Subsequent Performance

Peer personality might affect students' grades because peers affect students' human or social capital. If this is the case, peer effects may stay with the student and extend to outcomes beyond the course in which the peer-to-peer interaction takes place. Having two classrooms per period and multiple periods over time allows us to analyze the effects of peer personality in one classroom on performance in parallel and subsequent classrooms.

Table 9 provides estimates how peers in a given course affect students' performance in the parallel course and subsequent courses. Column (1) shows that the effect of having persistent peers on grades in parallel courses is positive but not statistically significant. When analyzing the effect of peer persistence on grades in subsequent courses in columns (2) and (3), we find that the point estimate is sizeable and statistically significant. Column (3) shows that a one standard deviation increase in peer persistence raises average future grades by 1.5 percent of a standard deviation. The size of this effect suggests a substantial nontransient effect of peer persistence on grades. This may occur because students accumulate human or social capital by interacting with persistent peers. They may acquire more skills or build a friendship network of persistent peers from which they benefit in current and future courses.

D. Heterogeneous Effects

From a policy perspective, it is important to ask how peer groups should be designed to maximize benefits arising from peer personality spillovers. To answer this question, we investigate whether the impact of peer personality is heterogeneous, depending on the students' own levels of the respective trait. Based on the students' own trait measures, we categorize students as having low

(P_i^l), medium (P_i^m), or high (P_i^h) levels of a particular trait depending on the tertile to which they belong.[24] We then interact students' own trait type with the peer personality measure, and estimate:

$$A_{icg} = \overline{PP}_{g-i} * (P_i^l\, \alpha_1{}' + P_i^m\, \alpha_2{}' + P_i^h \alpha_3{}')$$
$$+P_i^l\, \beta_4{}' + P_i^m\, \beta_5{}' + P_i^h\, \beta_6{}' + X_i \gamma' + \rho_{ic} + u_{ic}. \qquad (6)$$

Table 10 reports the estimates of this model, which allows for heterogeneous effects. Column (1) shows that students in the bottom and the top tertiles of the persistence distribution particularly benefit from a group of peers with high persistence. Students with medium persistence are not significantly affected when they are exposed to more-persistent peers. Column (2) shows that the results are again robust to the inclusion of other peer characteristics.[25] Columns (3) and (4) reveal that the effects that persist into future courses are driven by the same subgroup of students that cause the contemporaneous effects. These effects suggest that grouping high- and low-persistence students together could be an effective policy to increase overall achievement while reducing inequality in academic achievement.

We next turn to the question of whether peers who score very high or very low on specific personality measures influence performance in the same way as average peer personality in the group. For instance, having very persistent peers (shining lights) may be particularly beneficial for performance. Or, peers who have very limited persistence (bad apples) may be particularly detrimental for performance. We test for such nonlinearities by running regressions replacing the linear-in-means specification by the proportion of peers that belong to the top and bottom 10

[24] Tertiles are defined based on the global distribution and thus not at the section level.
[25] We also tested whether peer effects are heterogeneous by GPA of the student. Results are reported in Appendix Table A10. Students with a high GPA and those with a low GPA benefit most from persistent peers. Having risk-tolerant peers has the largest negative effect on performance among students from the bottom GPA-tertile.

percent of the overall distribution. Table 11 shows that students significantly benefit from students in the top 10 percent of the persistence distribution while they do not suffer from peers at the bottom 10 percent of the persistence distribution. This finding suggests that mixing low- and high-persistence students together could lead to improved average performance. We find also that students benefit from exposure to peers who are very unwilling to take risks.

VII. Conclusion

Previous literature on peer effects has studied the extent to which student performance depends on fellow students' achievement, gender, and race. This paper focuses on a different aspect of student interaction and shows that peer personality affects student achievement in university. In order to identify the causal impact of peers, we exploit the random assignment of students to university teaching sections.

Our results show that students who are exposed to more-persistent peers achieve higher grades. Peers' risk tolerance, self-confidence or anxiety do not significantly affect performance. We provide evidence that these personality peer effects are distinct from achievement peer effects. We also study the heterogeneity of personality peer effects and find that both students with high and low persistence benefit most from having highly persistent peers. Moreover, we investigate possible underlying mechanisms and find no evidence that students change their self-reported study hours when exposed to persistent peers. Students' evaluations of teachers also do not appear to be affected by persistent peers. We do, however, find suggestive evidence for complementarity between teacher and peer quality. Importantly, we also find that the effect of persistent peers is nontransient: exposure to persistent peers also increases grades in future periods. This suggests a human or social capital effect: Students who meet persistent peers at the beginning of their studies may develop better learning habits and a different network of friends that yields academic returns.

An inherent concern in the literature on personality traits is that other correlated factors may drive observed relationships. We believe that even if there are unobserved variables correlated with personality, it is still important to know that measures of peer personality are picking up a trait that creates spillovers. From a policy perspective, knowing that exposure to persistent peers will create learning spillovers matters. It is not necessarily important to know which traits correlated with persistence drive the result. As long as peer personality has predictive power for social spillovers, policies can be based on it.

Our results therefore have important implications for policies related to student tracking and the design of peer groups. They suggest that mixing low- and high-persistence students in study groups could potentially increase overall academic achievement. The results documented in this paper also have three important implications for the design of interventions and education policies that aim to improve socio-emotional skills. First, in settings in which treated and nontreated students interact, changes in peer personality may positively affect the educational attainment of nontreated students, which will make it more difficult to detect an intervention impact. Second, and more generally, the social returns of any intervention that enhances socio-emotional skills will be underestimated if positive spillovers of personality on other individuals outside the studied environment are neglected. Finally, if peers with a certain personality affect performance through particular behavior, then policy could also be targeted at stimulating that behavior directly. Future research could collect detailed behavioral measures to shed light on the mechanisms of personality peer effects.

REFERENCES

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M. (2017). When should you adjust standard errors for clustering? *NBER Working Paper 24003*.

Åkerlund, D., Golsteyn, B., Grönqvist, H., Lindahl, L. (2016). Time discounting and criminal behavior. *Proceedings of the National Academy of Sciences of the United States of America* 113(22), 6160-6165.

Almlund, M., Duckworth, A., Heckman, J., Kautz, T. (2011). Personality psychology and economics. In Handbook of the Economics of Education, Vol. 4, E. Hanushek, S. Machin, L. Woessman, eds. Amsterdam: Elsevier. 1-181.

Altonji, J.G., Elder, T.E., Taber, C.R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* 113(1), 151-184.

Angrist, J. D., Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *The American Economic Review* 94(5), 1613-1634.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

Booij, A. S., Leuven, E., Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *Review of Economic Studies* 84, 547-578.

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

Borghans, L., Golsteyn, B., Heckman, J. Humphries, J. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences of the United States of America* 113(47), 13354-13359.

Carrell, S. E., Fullerton, R. L., West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics* 27(3), 439-464.

Carrell, S. E., Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics* 2(1), 211-228.

Carrell, S. E., Hoekstra, M., Kuka, E. (2018). The long run effects of disruptive peers. *American Economic Review* 108(11)

Carrell, S. E., Sacerdote, B. I., West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica* 81(3), 855-882.

Chetty, R., Friedman, J. N., Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review* 104(9), 2633-2679.

De Giorgi, G., Pellizzari, M. (2013). Understanding social interactions: Evidence from the class room. *Economic Journal* 124, 917-953.

Dohmen, T., Falk, A, Huffman, D. Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioral outcomes. *Economic Journal* 119(536), 592-612.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G. (2011). Individual risk attitudes: measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9(3), 522-550.

Donald, D.G., and Lang, K. (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics* 89(2), 221-223.

Duckworth, A. L., Peterson, C., Matthews, M. D., Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology* 92(6), 1087-1101.

Duflo, E., Dupas, P., Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review* 101, 1739-1774.

Einiö, E. d'Este, R. (2018). The impact of Asian peers on scholastic achievement: Evidence from fertility shocks in the dragon year. *IZA Discussion Paper No* 11682.

Elsner, B, Isphording, I., Zölitz U. (2018). Achievement rank affects performance and major choices in college. *University of Zurich Discussion Paper No 270.*

Falk, A., Becker, A., Dohmen, T., Enke, B., Sunde, U., Huffman, D. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics* 133(4), 1645-1692.

Feld, J., Salamanca, N., Hamermesh, D. S. (2016). Endophilia or exophobia: Beyond discrimination. *The Economic Journal* 594, 1503-1527.

Feld, J., Salamanca, N., Zölitz, U. (forthcoming). Are professors worth It? The value-added and costs of university instructors with. *Journal of Human Resources*, forthcoming.

Feld, J., Zölitz, U. (2017). Understanding peer effects - on the nature, estimation and channels of peer effects. *Journal of Labor Economics* 35(2), 387-428.

Figlio, D. N. (2007). Boys named Sue: Disruptive children and their peers. *Education* 2(4), 376-394.

Golsteyn, B. H., Grönqvist, H., Lindahl, L. (2014). Adolescent time preferences predict lifetime outcomes. *The Economic Journal* 124(580), F739-F761.

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

Holland, B. S., Copenhaver, M. D. (1988). Improved Bonferroni-type multiple testing procedures. *Psychological Bulletin* 104(1), 145.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65-70.

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

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Paper* No. w7867.

Lavy, V., Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics* 3(2), 1-33.

Lyle, D. S. (2009). The effects of peer group heterogeneity on the production of human capital at West Point. *American Economic Journal: Applied Economics* 1, 69–84.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20, 141-151.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60, 531-540.

Martin, A. (2009). Motivation and engagement across the academic life span: A developmental construct validity study of elementary school, high school, and university/college students. *Educational and Psychological Measurement* 69(5), 794-824.

Martin, A.J. (2011). Courage in the classroom: Exploring a new framework predicting academic performance and engagement. *School Psychology Quarterly*, 26, 145-160.

Oosterbeek, H., van Ewijk, R. (2014). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review* 38, 51-63.

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. Hanushek, S. Machin, L. Woessmann (eds.), *Handbook of the Economics of Education*, Elsevier, edition 1, volume 3, number 3, 249-277.

Shure, N. (2017). Non-cognitive peer effects in secondary education. Unpublished manuscript.

Sidak, Z. (1968). On multivariate normal probabilities of rectangles: Their dependence on correlations. *Ann. Math. Statist*. 39, 1425-34.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73(3), 751-754.

Stepner, M. (2013). VAM: Stata module to compute teacher value-added measures, *Statistical Software Components* S457711, Boston College Department of Economics.

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., Martinsson, P. (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association* 13(3), 421-452.

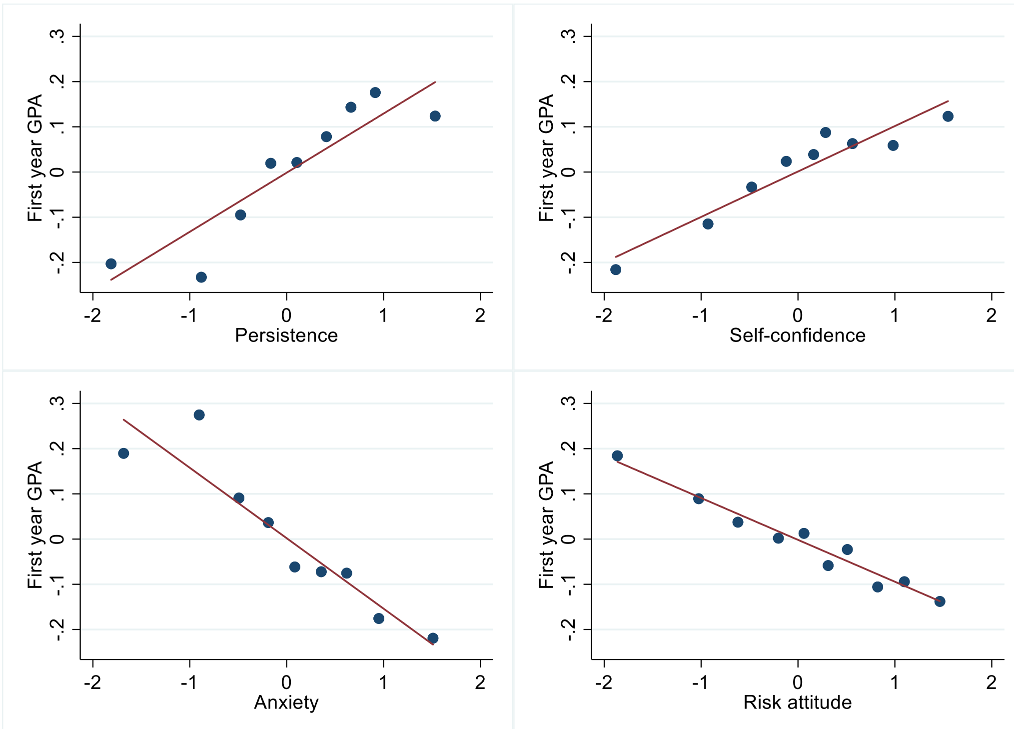
Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: evidence from a randomized experiment. *American Economic Review* 95(2), 199-203.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: evidence from a natural experiment. *Review of Economics and Statistics* 85(1), 9-23.

Zölitz, U., Feld, J. (2019). The effect of peer gender on major choice. *University of Zurich Discussion Paper No 270.*

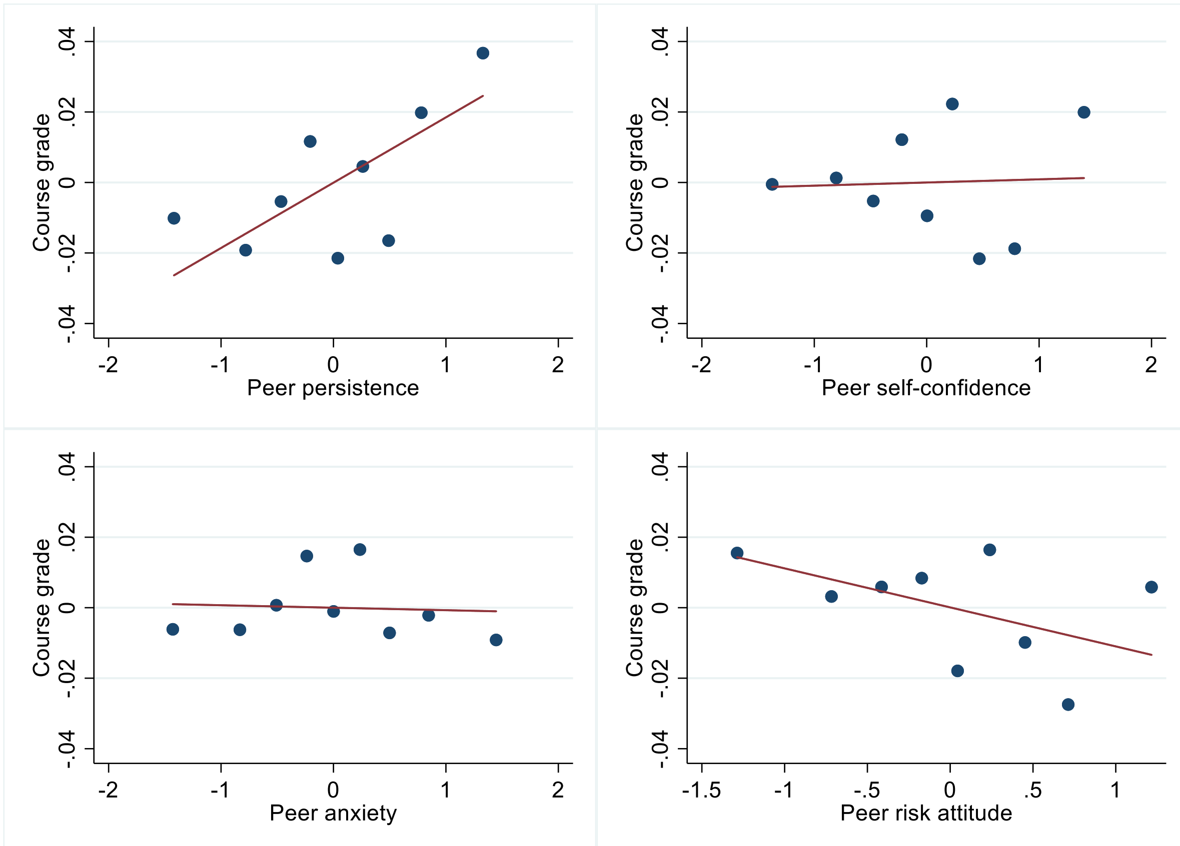
FIGURES AND TABLES

Figure 1: Student Personality and Performance



NOTE.— These plots visualize the regression results reported in column (6) of Table 4. The construction of these binned scatter plots follows Chetty, Friedman, and Rockoff (2014). We first regress course grades on the set of controls included in column (6) of Table 4 to obtain the residualized course grades. Next, we rank-order our measures of personality and split them into equally sized bins. We then plot the mean of the residualized course grades within each bin against the normalized mean value of personality in that bin. N = 4,383.

Figure 2: Impact of Peer Personality on Performance



NOTE.—These plots visualize the regression results reported in column (6) of Table 5. The construction of these binned scatter plots follows Chetty, Friedman, and Rockoff (2014). We first regress course grades on the set of controls included in column (6) of Table 5 to obtain the residualized course grades. Next, we rank-order our measures of peer personality and split them into equally sized bins. We then plot the mean of the residualized course grades within each bin against the normalized mean value of peer personality in that bin. N = 16,155.

Table 1: Descriptive Statistics

Panel A: Data overview	(1) N
Number of students	4,383
Number of student course registrations	17,512
Number of student course grades	16,155
Number of total sections (tutorials)	1,357
Number of courses (course-year observations)	58
Number of courses per student	3.995
Number of grades per student	3.686
Average number of students per section (class size)	12.904
Average number of sections per course	23.397

Panel B: Student-level characteristics	(1) N	(2) mean	(3) sd	(4) min	(5) max
Female	4,383	0.41	0.49	0.00	1.00
Dutch	4,383	0.24	0.43	0.00	1.00
German	4,383	0.47	0.50	0.00	1.00
GPA	4,383	6.66	1.60	1.50	10.00
High school math major	4,383	0.34	0.47	0.00	1.00
Risk attitude	4,383	4.74	1.31	1.00	7.00
Persistence	4,383	5.38	0.82	1.75	7.00
Self-confidence	4,383	5.81	0.70	2.50	7.00
Anxiety	4,383	4.66	1.22	1.00	7.00

Panel C: Section-level peer characteristics	(1) N	(2) mean	(3) sd	(4) min	(5) max
Peer persistence	1,357	5.36	0.25	4.50	6.75
Peer self-confidence	1,357	5.82	0.22	4.75	7.00
Peer anxiety	1,357	4.64	0.38	1.50	5.63
Peer risk attitude	1,357	4.78	0.49	3.00	7.00
Peer GPA	1,357	6.55	0.64	2.33	8.62
Proportion of peers with high school math major	1,357	0.35	0.15	0.00	1.00

NOTE.—Estimation sample summary statistics.

Table 1: Descriptive Statistics (continued)

Panel D: Student performance, study effort, and course evaluations	(1) N	(2) mean	(3) sd	(4) min	(5) max
Course dropout	17,512	0.08	0.27	0.00	1.00
Course grade	16,155	6.49	1.68	0.50	10.00
Self-reported study hours	8,656	13.41	8.44	0.00	60.00
The tutor sufficiently mastered the course content	9,339	4.39	0.89	1.00	5.00
The tutor stimulated the transfer of what I learned in this course to other contexts	9,330	4.02	1.04	1.00	5.00
The tutor was enthusiastic in guiding our group	9,327	4.12	1.07	1.00	5.00
Evaluate the overall functioning of your tutor in this course with a grade	9,140	7.92	2.00	1.00	10.00
Working in tutorial groups with my fellow students helped me to better understand the subject matter	9,337	4.07	0.87	1.00	5.00
My tutorial group has functioned well.	9,315	4.02	0.89	1.00	5.00
Please give an overall grade for the quality of this course	9,103	7.20	1.74	1.00	10.00

NOTE.—Estimation sample summary statistics.

Table 2: Measurement of Student Personality

Trait	Definition	Measurement
Persistence	How much students keep trying to work out an answer or to understand a problem even when that problem is difficult or is challenging.	1. "If I can't understand my university work at first, I keep going over it until I do" 2. "If my homework is difficult, I keep working at it trying to figure it out" 3. "When I'm taught something that doesn't make sense, I spend time to try to understand it" 4. "I'll keep working at difficult university work until I think I've worked it out"
Self-confidence	Students' belief and confidence in their ability to understand or to do well in their studies.	Four items, e.g. "If I try hard, I believe I can do my university work well."
Anxiety	Feeling nervous when thinking about their studies and worrying about not doing well in their studies.	1. "When exams and assignments are coming up, I worry a lot" 2. "I worry about failing exams and assignments" 3. "When I do tests or exams I don't feel very well" 4. "In terms of my university work, I'd call myself a worrier"
Risk attitude	Willingness to take risks.	"In general, how willing are you to take risks?"

NOTE.—All concepts are measured on a scale of 1 to 7. Persistence, Self-confidence, and Anxiety were taken from the Student Motivation Scale (Martin 2009). The items for self-confidence are not publicly available. The persistence and anxiety items have been made publicly available in Martin (2011). For a discussion and validation of the measure of risk attitude, see Dohmen et al. (2011).

Table 3: Test for Random Assignment

	(1)	(2)	(3)	(4)	(5)
	Peer persistence	Peer self-confidence	Peer anxiety	Peer risk attitude	Prop peers with high school math major
Persistence	0.0108				
	(0.015)				
Self-confidence		0.0088			
		(0.009)			
Anxiety			-0.0149		
			(0.011)		
Risk attitude				0.0034	
				(0.010)	
High school math major					0.0004
					(0.003)
Observations	17,512	17,512	17,512	17,512	17,512
R-squared	0.169	0.136	0.242	0.450	0.180

NOTE.—The dependent variable in all columns is the standardized section-level leave-out mean of the respective personality characteristic, i.e. the average peer persistence in a section excluding the student's own personality. All models are estimated with ordinary least squares regressions that include course-times-year fixed effects and controls for gender and nationality (Dutch, German). Following the Guryan, Kroft, and Notowidigdo (2009) correction method, we control for the course-level leave-out mean in all estimations. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Student Personality and Student Achievement

	(1) Std first year GPA	(2) Std first year GPA	(3) Std first year GPA	(4) Std first year GPA	(5) Std first year GPA	(6) Std first year GPA
Persistence	0.1307***					0.1058***
	(0.015)					(0.017)
Self-confidence		0.1005***				0.0354**
		(0.014)				(0.016)
Anxiety			-0.1558***			-0.1427***
			(0.015)			(0.015)
Risk attitude				-0.0926***		-0.1098***
				(0.015)		(0.015)
High school math major					0.2524***	0.2159***
					(0.031)	(0.030)
Observations	4,383	4,383	4,383	4,383	4,383	4,383
R-squared	0.092	0.086	0.098	0.084	0.090	0.134

NOTE.— This table shows student-level regressions with one observation representing one individual. The dependent variable is the end of the first year standardized GPA based on all first year courses. All models are estimated with ordinary least squares regressions that include cohort-times-study-program fixed effects and that control for gender and nationality (Dutch, German). Robust standard errors are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 5: The Effect of Peer Personality on Student Performance

	(1) Std. grade	(2) Std. grade	(3) Std. grade	(4) Std. grade	(5) Std. grade	(6) Std. grade	(7) Std. grade
Peer persistence	0.0193***				0.0186***	0.0180***	0.0175***
	(0.005)				(0.006)	(0.006)	(0.006)
Peer self-confidence		0.0085			0.0008	0.0010	0.0011
		(0.006)			(0.007)	(0.007)	(0.006)
Peer anxiety			-0.0015		-0.0011	-0.0012	-0.0023
			(0.005)		(0.005)	(0.005)	(0.005)
Peer risk attitude				-0.0111**	-0.0114**	-0.0116**	-0.0109*
				(0.005)	(0.005)	(0.006)	(0.006)
Observations	16,155	16,155	16,155	16,155	16,155	16,155	16,155
R-squared	0.640	0.640	0.640	0.640	0.641	0.644	0.644
P-value joint significance of peer personality variables	-	-	-	-	.0012	.0017	.0043
Controlling for peer GPA, gender, nationality, and high school math major	NO	NO	NO	NO	NO	NO	YES
Indicators for scheduling conflicts	NO	NO	NO	NO	NO	YES	YES

NOTE.—The dependent variable is the standardized course grade. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as measures for the respective own personality trait as well as gender, nationality (Dutch, German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. The estimation sample of this table consists of all student-course observations with nonmissing grades. The number of observations is lower than in Table 3 because some students drop out of the course. Appendix Table A2 shows that peer personality does not affect drop out. Table A3 shows the randomization check for this estimation sample. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Peer Personality and Overall First-Year Performance

	(1) Number of first-year courses passed	(2) Successfully completed the first year	(3) Std. GPA at end of first year
Peer persistence	0.0433*	0.0131*	0.0257**
	(0.024)	(0.007)	(0.011)
Peer self-confidence	0.0473**	0.0028	0.0023
	(0.023)	(0.007)	(0.011)
Peer anxiety	0.0211	0.0011	0.0045
	(0.024)	(0.007)	(0.011)
Peer risk attitude	-0.0128	-0.0074	-0.0076
	(0.032)	(0.009)	(0.015)
Observations	4,082	4,082	3,931
R-squared	0.508	0.451	0.723
Mean dependent variable	2.8 out of 6	0.48	0
Program-times-year fixed effects	YES	YES	YES

NOTE.— This table shows student-level regressions in which one observation represents one individual. The dependent variable in column (1) is the total number of courses that a student passed during the first study year. The dependent variable in column (2) is an indicator for whether a student successfully completed the first year based on the institution's requirements. The dependent variable in column (3) is a student's standardized GPA based on all first-year courses. All models are estimated with ordinary least squares regressions that include program-year fixed effects, measures for the respective own personality traits and own characteristics such as gender, nationality (Dutch and German), and the pretreatment GPA measured in the first period of the first year. In this table, the unit of observation is the student. Robust standard errors are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 7: Mechanisms I – Student Course Evaluations

	(1) Study hours	(2) Instructor quality index	(3) Peer interaction index	(4) Overall course quality
Peer persistence	0.0238	-0.0166	-0.0066	-0.0020
	(0.111)	(0.022)	(0.018)	(0.017)
Peer self-confidence	0.0186	-0.0053	-0.0029	-0.0172
	(0.103)	(0.025)	(0.018)	(0.017)
Peer anxiety	0.1566	-0.0566***	-0.0296*	-0.0605***
	(0.116)	(0.018)	(0.016)	(0.016)
Peer risk attitude	-0.0893	0.0026	0.0098	0.0044
	(0.107)	(0.022)	(0.021)	(0.019)
Observations	8,463	9,170	9,143	8,894
R-squared	0.134	0.174	0.073	0.181
Controlling for peer gender nationality, high school math major, and peer GPA	YES	YES	YES	YES
Indicators for scheduling conflicts	YES	YES	YES	YES

NOTE.—The dependent variable in column (1) is student's self-reported weekly study hours. The dependent variable in column (2) is the instructor quality index defined as the average of the standardized value of the four evaluation items: "Evaluate the overall functioning of your tutor in this course with a grade," "The tutor sufficiently mastered the course content," "The tutor stimulated the transfer of what I learned in this course to other contexts," and "The tutor was enthusiastic in guiding our group." The dependent variable in column (3) is the peer interaction index, defined as the average of the standardized value of the two evaluation items: "My tutorial group has functioned well" and "My fellow students helped me to better understand the subject matter." The dependent variable in column (4) is based on the question "Please give an overall grade for the quality of this course." All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch, German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. The estimation sample of this table consists of all student-course observations with nonmissing values of the respective evaluation item(s). Appendix Table A7 shows that peer personality does not affect the survey response probability. Table A8 shows that our main results are very similar for the subsample included in this regression table. Robust standard errors clustered at the course level are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 8: Mechanisms II – Heterogeneous Effects by Teacher Quality

	(1)	(2)	(3)
Dependent Variable:		Std. Course Grade	
Teacher Value-Added Subsample:	High-quality teacher	Medium-quality teacher	Low-quality teacher
Peer persistence	0.0416***	-0.0175	-0.0074
	(0.014)	(0.017)	(0.012)
Peer self-confidence	-0.0104	0.0436***	0.0011
	(0.012)	(0.013)	(0.011)
Peer anxiety	-0.0062	0.0011	0.0055
	(0.012)	(0.011)	(0.011)
Peer risk attitude	-0.0148	0.0025	-0.0174
	(0.017)	(0.015)	(0.014)
Observations	3,583	3,252	3,556
R-squared	0.643	0.654	0.620
Controlling for peer gender nationality, high school math major, and peer GPA	YES	YES	YES
Indicators for scheduling conflicts	YES	YES	YES

NOTE.—The dependent variable is the standardized course grade. Teacher quality is based on the respective tertile of teacher value-added of the classroom instructor. Teacher value-added is constructed using the VAM program by Stepner (2013) that follows Chetty et al. (2014). All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch, German), third-order polynomial in GPA, course-year fixed effects, class size, and the number of students with nonmissing personality measures. The estimation sample of this table consists of all student-course observations with nonmissing teacher value-added. The computation of teacher value-added requires observing instructors teaching the same course in multiple years, which leads to missing teacher value-added measures for 5,764 student-course observations. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: Performance in Parallel and Subsequent Courses

	(1) Std. grade parallel course	(2) Std. grade next period	(3) Std. average future grades
Peer persistence	0.0071	0.0117*	0.0149***
	(0.010)	(0.006)	(0.005)
Peer self-confidence	-0.0036	-0.0041	-0.0017
	(0.010)	(0.008)	(0.007)
Peer anxiety	0.0146	0.0007	-0.0064
	(0.011)	(0.007)	(0.007)
Peer risk attitude	0.0035	0.0000	0.0053
	(0.010)	(0.007)	(0.007)
Observations	7,030	9,422	9,227
R-squared	0.567	0.559	0.572
Controlling for peer gender, nationality, high school math major, and peer GPA	YES	YES	YES
Indicators for scheduling conflicts	YES	YES	YES

NOTE.— The table provides an analysis of how peer personality affects performance in parallel and subsequent courses. The dependent variable in column (1) is the standardized course grade of the parallel course taken at the same time. The dependent variable in column (2) is the standardized average course grade obtained in the subsequent period. The dependent variable in column (3) is the average standardized course grade obtained in all subsequent periods. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch, German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. The estimation sample excludes periods 1 and 2 in which peer groups in the course and parallel course are identical. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 10: Effect of Peer Personality on Student Performance – Heterogeneous Effects

	(1) Std. grade	(2) Std. grade	(3) Average future grades	(4) Average future grades
Low persistence * Peer persistence	0.0269***	0.0265***	0.0185**	0.0188**
	(0.009)	(0.009)	(0.008)	(0.009)
Medium persistence * Peer persistence	0.0098	0.0093	-0.0004	0.0003
	(0.008)	(0.008)	(0.009)	(0.009)
High persistence * Peer persistence	0.0159*	0.0153*	0.0253**	0.0258**
	(0.009)	(0.009)	(0.011)	(0.011)
Low self-confidence * Peer self-confidence	0.0024	0.0026	0.0009	0.0018
	(0.009)	(0.009)	(0.009)	(0.010)
Medium self-confidence * Peer self-confidence	0.0074	0.0078	0.0026	0.0026
	(0.012)	(0.012)	(0.010)	(0.010)
High self-confidence * Peer self-confidence	-0.0055	-0.0051	-0.0067	-0.0061
	(0.010)	(0.009)	(0.012)	(0.011)
Low anxiety * Peer anxiety	-0.0105	-0.0117	-0.0148	-0.0139
	(0.008)	(0.008)	(0.012)	(0.012)
Medium anxiety * Peer anxiety	0.0015	0.0003	-0.0251***	-0.0248***
	(0.010)	(0.011)	(0.009)	(0.009)
High anxiety * Peer anxiety	0.0080	0.0068	0.0214**	0.0221**
	(0.008)	(0.008)	(0.009)	(0.009)
Low risk * Peer risk attitude	0.0031	0.0034	0.0112	0.0100
	(0.009)	(0.009)	(0.012)	(0.012)
Medium risk * Peer risk attitude	-0.0111	-0.0108	0.0083	0.0073
	(0.008)	(0.008)	(0.008)	(0.008)
High risk * Peer risk attitude	-0.0260**	-0.0258**	0.0009	0.0009
	(0.011)	(0.011)	(0.011)	(0.011)
Observations	16,155	16,155	9,227	9,227
R-squared	0.644	0.644	0.573	0.574
P-value joint significance of peer variables	.0014	.005	<.001	<.001
Controlling for peer gender, nationality, GPA, and high school math major	NO	YES	NO	YES
Indicators for scheduling conflicts	YES	YES	YES	YES

NOTE.—The dependent variable in columns (1) and (2) is the standardized course grade. The dependent variable in columns (3) and (4) is the standardized average future course grades. All models are estimated with ordinary least squares regressions that include dummies for the respective own personality trait tertile, gender, nationality (Dutch, German), third-order polynomial in GPA, course-year fixed effects, class size, and the number of students with nonmissing personality measures. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 11: Peer Personality – Bad Apples and Shining Lights

	(1) Std. grade	(2) Std. grade	(3) Std. grade	(4) Std. grade	(5) Std. grade	(6) Std. grade	(7) Average future grades
Proportion of top 10% persistent peers	0.1047**				0.0945*	0.0889*	0.0401
	(0.049)				(0.057)	(0.053)	(0.065)
Proportion of bottom 10% persistent peers	-0.0431				-0.0448	-0.0422	-0.0458
	(0.059)				(0.058)	(0.058)	(0.064)
Proportion of top 10% self-confident peers		0.0445			-0.0053	-0.0038	-0.0428
		(0.051)			(0.056)	(0.057)	(0.073)
Proportion of bottom 10% self-confident peers		-0.0228			0.0103	0.0095	-0.0417
		(0.059)			(0.060)	(0.052)	(0.060)
Proportion of top 10% anxious peers			0.0058		0.0061	-0.0028	-0.0839
			(0.056)		(0.054)	(0.054)	(0.064)
Proportion of bottom 10% anxious peers			0.0724		0.0708	0.0754	0.0190
			(0.050)		(0.047)	(0.056)	(0.068)
Proportion of top 10% risk-tolerant peers				0.0232	0.0094	0.0142	0.0618
				(0.042)	(0.041)	(0.041)	(0.047)
Proportion of bottom 10% risk-tolerant peers				0.1428***	0.1323**	0.1304**	0.0405
				(0.054)	(0.054)	(0.051)	(0.056)
Observations	16,155	16,155	16,155	16,155	16,155	16,155	9,227
R-squared	0.644	0.644	0.644	0.644	0.644	0.644	0.572
Controlling for peer gender, nationality, GPA, and high school math major	NO	NO	NO	NO	NO	YES	YES
Indicators for scheduling conflicts	YES	YES	YES	YES	YES	YES	YES

NOTE.—The dependent variable in columns (1)–(6) is the standardized course grade. The dependent variable in Column (7) is the standardized average future course grade. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Additional controls include indicators for own top and bottom 10 percent status for each of the four traits. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

APPENDIX A

Table A1: Pairwise Correlations between Measures of Student Personality

	(1) Persistence	(2) Self-confidence	(3) Anxiety	(4) Risk attitude
Persistence	1			
Self-confidence	0.4579***	1		
	(0.014)			
Anxiety	-0.0268*	-0.1842***	1	
	(0.016)	(0.015)		
Risk attitude	0.0202	0.1068***	-0.1299***	1
	(0.016)	(0.015)	(0.016)	

NOTE.—The table shows student-level pairwise correlations between the measures of student personality we use throughout the paper. N = 4,383. Standard errors are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A2: Peer Personality and Course Dropout

	(1) Course dropout	(2) Course dropout	(3) Course dropout	(4) Course dropout	(5) Course dropout	(6) Course dropout	(7) Course dropout
Peer persistence	0.0017				0.0015	0.0015	0.0006
	(0.002)				(0.003)	(0.003)	(0.003)
Peer self-confidence		0.0005			-0.0005	0.0001	0.0004
		(0.002)			(0.002)	(0.002)	(0.002)
Peer anxiety			-0.0007		-0.0001	0.0001	-0.0007
			(0.002)		(0.002)	(0.002)	(0.002)
Peer risk attitude				-0.0042*	-0.0035	-0.0037	-0.0033
				(0.002)	(0.003)	(0.003)	(0.003)
Observations	17,512	17,512	17,512	17,512	17,512	17,512	17,512
R-squared	0.176	0.176	0.175	0.175	0.177	0.179	0.179
Controlling for peer gender nationality, high school math major, and peer GPA	NO	NO	NO	NO	NO	NO	YES
Indicators for scheduling conflicts	NO	NO	NO	NO	NO	YES	YES

NOTE.—The dependent variable is an indicator for whether the student dropped out of the course. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A3: Randomization Check for Sample with Nonmissing Grades

	(1) Peer persistence	(2) Peer self-confidence	(3) Peer anxiety	(4) Peer risk attitude	(5) Prop peers with high school math major
Persistence	0.0055				
	(0.016)				
Self-confidence		-0.0018			
		(0.012)			
Anxiety			-0.0133		
			(0.012)		
Risk attitude				-0.0019	
				(0.013)	
High school math major					0.0031
					(0.003)
Observations	16,155	16,155	16,155	16,155	16,155
R-squared	0.167	0.131	0.242	0.447	0.166

NOTE.—The dependent variable in all columns is the standardized section-level leave-out mean of the respective personality characteristic, i.e. the average peer persistence in a section excluding the student's own personality. All models are estimated with ordinary least squares regressions that include course-times-year fixed effects and controls for gender and nationality (Dutch, German). Following the Guryan, Kroft, and Notowidigdo (2009) correction method, we control for the course-level leave-out mean in all estimations. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A4: Robustness to Including Additional Measures of Cognitive Ability

	(1) Std. grade	(2) Std. grade	(3) Std. grade	(4) Std. grade	(5) Std. grade	(6) Std. grade
Peer persistence	0.0179***	0.0175***	0.0175***	0.0174***	0.0172***	0.0171***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Peer self-confidence	0.0010	0.0012	0.0011	-0.0002	0.0017	0.0010
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Peer anxiety	-0.0010	-0.0017	-0.0023	-0.0020	-0.0028	-0.0020
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Peer risk attitude	-0.0114**	-0.0110**	-0.0109*	-0.0103*	-0.0099*	-0.0100*
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	16,155	16,155	16,155	16,032	16,039	15,982
R-squared	0.644	0.644	0.644	0.644	0.645	0.645
Indicators for scheduling conflicts	YES	YES	YES	YES	YES	YES
Controlling for peer gender and nationality	NO	YES	YES	YES	YES	YES
Controlling for peer high school math major and GPA	NO	YES	YES	YES	YES	YES
Controlling for peer math entry test score	NO	NO	NO	YES	NO	YES
Controlling for peer statistics entry test score	NO	NO	NO	NO	YES	YES

NOTE.—The dependent variable is the standardized course grade. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch and German), course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. In columns (2) to (6) we include third-order polynomial in GPA and a dummy indicating whether the student was a math major. The estimation sample of this table consists of all student-course observations with nonmissing grades. The number of observations is lower in columns (4)–(6) due to missing values of math and statistics test scores. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A5: Corrections for Multiple Hypothesis Testing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Coefficient	Standard error	Original uncorrected p-value from main results table	p-value for one-step method following		p-value for step-down method following		p-value for step-up method following	
Correction method:				Bonferroni correction	Sidak (1968) correction	Holm (1979)	Holland & Copenhaver (1988)	Benjamini & Hochberg (1995)	Simes (1986)
Peer persistence	0.0175	0.0063	0.0068	0.0272	0.0270	0.0272	0.0270	0.0272	0.0272
Peer self-confidence	0.0011	0.0065	0.8636	1.0000	0.9997	1.0000	0.8772	0.8636	0.8636
Peer anxiety	-0.0023	0.0051	0.6495	1.0000	0.9849	1.0000	0.8772	0.8636	0.8636
Peer risk attitude	-0.0109	0.0056	0.0544	0.2177	0.2006	0.1633	0.1545	0.1633	0.1088

NOTE.— Table reports estimated p-values of the respective peer personality coefficients. Columns (1) to (3) are taken from our main specification in Table 5, column (7). Regressions in columns (4)–(9) follow the same specification.

Table A6: Main Results with Standard Errors Clustered at the Section Level

	(1) Std. grade	(2) Std. grade	(3) Std. grade	(4) Std. grade	(5) Std. grade	(6) Std. grade	(7) Std. grade
Peer persistence	0.0193***				0.0186***	0.0180***	0.0175***
	(0.005)				(0.006)	(0.006)	(0.006)
Peer self-confidence		0.0085*			0.0008	0.0010	0.0011
		(0.005)			(0.006)	(0.006)	(0.006)
Peer anxiety			-0.0015		-0.0011	-0.0012	-0.0023
			(0.005)		(0.005)	(0.005)	(0.006)
Peer risk attitude				-0.0111*	-0.0114*	-0.0116*	-0.0109*
				(0.006)	(0.006)	(0.006)	(0.006)
Observations	16,155	16,155	16,155	16,155	16,155	16,155	16,155
R-squared	0.640	0.640	0.640	0.640	0.641	0.644	0.644
Controlling for peer gender, nationality, and high school math major	NO	NO	NO	NO	NO	NO	YES
Indicators for scheduling conflicts	NO	NO	NO	NO	NO	YES	YES

NOTE.—The dependent variable is the standardized course grade. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Robust standard errors clustered at the *section* level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A7: Peer Personality and Course Evaluation Survey Response

	(1) Course evaluation response
Peer persistence	-0.0051
	(0.005)
Peer self-confidence	0.0013
	(0.005)
Peer anxiety	0.0058
	(0.005)
Peer risk attitude	-0.0052
	(0.006)
Observations	17,512
R-squared	0.173
Mean dependent variable	0.4943
P-value joint significance of peer variables	.6944

NOTE.—The dependent variable is an indicator for students' participation in the course evaluations. The response rate is 49 percent. The model is estimated with ordinary least squares regressions that include measures for the respective own personality trait as well gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A8: Main Results for Students who Participated in Student Course Evaluations

	(1) Std. grade	(2) Std. grade	(3) Std. grade	(4) Std. grade	(5) Std. grade	(6) Std. grade	(7) Std. grade	(8) Std. grade
Peer persistence	0.0191***					0.0195**	0.0187**	0.0186**
	(0.007)					(0.008)	(0.008)	(0.008)
Peer self-confidence		0.0063				-0.0021	-0.0008	-0.0005
		(0.008)				(0.010)	(0.010)	(0.010)
Peer anxiety			-0.0011			0.0007	0.0007	0.0011
			(0.008)			(0.007)	(0.007)	(0.007)
Peer risk attitude				-0.0101		-0.0088	-0.0087	-0.0081
				(0.007)		(0.007)	(0.007)	(0.008)
Peer GPA					0.0062	0.0033	0.0034	-0.0005
					(0.007)	(0.008)	(0.008)	(0.007)
Observations	8,463	8,463	8,463	8,463	8,463	8,463	8,463	8,463
R-squared	0.652	0.651	0.652	0.651	0.651	0.652	0.655	0.656
Controlling for peer gender, peer nationality, peer math major, and peer GPA	NO	NO	NO	NO	NO	NO	NO	YES
Indicators for scheduling conflicts	NO	NO	NO	NO	NO	NO	YES	YES

NOTE.— The dependent variable is the standardized course grade. All models are estimated with ordinary least squares regressions that include measures for the respective own personality trait as well as gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A9: Determinants of Self-Reported Study Hours

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Study hours	Study hours	Study hours	Study hours	Study hours
Persistence	1.2442***				1.5170***
	(0.102)				(0.101)
Self-confidence		0.1074			-0.4409***
		(0.094)			(0.082)
Anxiety			0.6410***		0.6616***
			(0.104)		(0.106)
Risk attitude				-0.2847***	-0.2537***
				(0.085)	(0.085)
Observations	8,463	8,463	8,463	8,463	8,463
R-squared	0.123	0.104	0.108	0.104	0.133

NOTE.—The dependent variable is students' self-reported weekly study hours. All models are estimated with ordinary least squares regressions that include gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Personality measures are standardized to mean zero and a standard deviation of one. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A10: Effect of Peer Personality on Student Performance – Heterogeneous Effects by Student GPA

	(1) Std. grade
Peer persistence * High GPA	0.0210**
	(0.009)
Peer persistence * Medium GPA	0.0023
	(0.009)
Peer persistence * Low GPA	0.0292***
	(0.011)
Peer self-confidence * High GPA	0.0035
	(0.011)
Peer self-confidence * Medium GPA	0.0074
	(0.009)
Peer self-confidence * Low GPA	-0.0071
	(0.013)
Peer anxiety * High GPA	-0.0013
	(0.008)
Peer anxiety * Medium GPA	0.0057
	(0.007)
Peer anxiety * Low GPA	-0.0114
	(0.010)
Peer risk attitude * High GPA	0.0023
	(0.015)
Peer risk attitude * Medium GPA	-0.0017
	(0.009)
Peer risk attitude * Low GPA	-0.0338**
	(0.013)
Observations	16,155
R-squared	0.644
Controlling for peer gender, nationality, and high school math major	YES
Indicators for scheduling conflicts	YES

NOTE.—The dependent variable is the standardized course grade. Coefficients are estimated with ordinary least squares regression that includes measures for the respective own personality trait as well as gender, nationality (Dutch and German), third-order polynomial in GPA, course-times-year fixed effects, class size, and the number of students with nonmissing personality measures. Additional controls include indicators for own low, medium, or high GPA status. Robust standard errors clustered at the course level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

APPENDIX B: Validation of Student Motivation Scale

This paper studies the effects of peer personality on grades. How reliable are the personality measures we use in this paper? How do our measures of personality relate to measures that are more widely used in the literature? To shed light on these two important questions, we conducted three additional surveys in 2018 among first-year economics students at the business school (N = 625).

A. Reliability

In order to understand how reliable our personality measures are, we analyze the test-retest correlation of the traits in the first two additional surveys. The first survey was taken in the first week of students' first course at the business school and includes the Student Motivation Scale—the personality measures we use throughout this paper. The second survey was conducted four weeks later in the fifth week of the first course and also includes the Student Motivation Scale. We ran this second survey to analyze the test-retest correlation of these traits.

Table B1 shows the test-retest correlation of the traits we study in this paper. Our results indicate that the test-retest correlation between persistence in week 1 and persistence in week 5 is 0.57, while the test-retest correlation of anxiety is 0.75.

B. Relationship between Student Motivation Scale, the Big Five, and Grit

How does our measure of persistence and other personality measures relate to measures that are used more often in the literature? To provide evidence on comparability to other measures, we conducted a third survey, three weeks after the second survey, which included measures of grit and Big Five personality traits, conscientiousness and neuroticism.[26]

Table B1 shows that the correlation between persistence in week 5 and Big Five conscientiousness is 0.41. This correlation is very high when compared to the test-retest correlation of persistence (0.57). The correlation between persistence in week 5 and grit is 0.43, which is similar to the correlation between conscientiousness and persistence. The correlation between our anxiety measure and neuroticism is 0.60. This correlation is also high when compared to the high test-retest correlation we find for anxiety (0.75).

Taken together, Table B1 shows that persistence is highly correlated with conscientiousness and grit. Are these measures merely correlated to each other or do they in fact represent the same concept? Although correlations are often used to validate measures, these correlations do not necessarily show the extent to which the two measures are, in fact, the same.

In order to probe deeper into the question of whether persistence is a facet of conscientiousness or a facet of grit and whether anxiety is a facet of neuroticism, we perform three factor analyses (see Table B2).

We first perform a factor analysis with all persistence and all conscientiousness items and find that the Eigenvalue of the first factor is 3.71 and that of the second factor is 1.29. Kaiser (1960) suggests that items measure the same trait as long as the Eigenvalue in the factor analysis is above one for one factor only. The results therefore indicate that although persistence is strongly correlated with conscientiousness, it is not a facet of conscientiousness. Our second factor analysis includes all persistence and all grit items. The Eigenvalue of the first factor is 3.25 and that of the

[26] Grit is measured with six items from the 12-item grit effort scale (Duckworth et al. 2007). Conscientiousness is measured with the NEO-domain 10-item IPIP scale. We chose to have a lag between the measures to reduce the likelihood that students remember their answers on the first set of questions, which would have led to consistency bias.

second factor is 0.92. We therefore conclude that persistence is a facet of grit. Our third factor analysis includes all anxiety and neuroticism items. The Eigenvalue of the first factor is 4.21 and that of the second factor is 0.74. We therefore conclude that anxiety is a facet of neuroticism.

In sum, the results indicate that our measure of persistence is in fact a measure of grit and that our measure of anxiety is a measure of neuroticism.

Table B1: Correlations between Martin Measures, Big Five Measures, and Grit

	(1) Persistence week 1	(2) Persistence week 5	(3) Anxiety week 1	(4) Anxiety week 5	(5) Grit week 8	(6) Big Five conscien-tiousness week 8	(7) Big Five neuroticism week 8
Persistence week 1	1						
Persistence week 5	0.5698***	1					
Anxiety week 1	0.012	0.0167	1				
Anxiety week 5	-0.049	-0.0355	0.7513***	1			
Grit week 8	0.3915***	0.4306***	-0.0599	-0.0174	1		
Big Five conscientiousness week 8	0.3303***	0.4060***	0.0427	0.0046	0.5725***	1	
Big Five neuroticism week 8	-0.0103	0.0527	0.6183***	0.6016***	-0.0607	0.0617	1

NOTE.— N = 617. All personality measures are collected during the first course period among first-year students of the 2018–2019 cohort using three student online surveys in week 1, week 5, and 8 of the term. Persistence and anxiety are measured with Martin's Student Motivation Scale during the first and fifth week of the course period. Grit, conscientiousness, and neuroticism are measured three weeks later. Grit is measured with 6 items from the 12-item grit effort scale (Duckworth et al. 2007). Conscientiousness is measured using the relevant 10 items from the NEO-domain 50-item IPIP scale. Neuroticism is measured using the three neuroticism items from the BFI-S 15 items scale. * p<0.1, ** p<0.05, *** p<0.01.

Table B2: Eigenvalues of Three Factor Analyses

	(1) Persistence & Big Five conscientiousness	(2) Persistence & grit effort	(3) Anxiety & Big Five neuroticism
Factor 1	3.71	3.23	4.21
Factor 2	1.29	0.92	0.74

NOTE.— N = 617. Each column reports the first two Eigenvalues that result from a different factor analysis. The factor analysis reported in column (1) includes all separate items of persistence and Big Five conscientiousness; the factor analysis in column (2) includes all separate items of persistence and grit effort; the factor analysis in column (3) includes all separate items of anxiety and Big Five neuroticism. Persistence is measured with Martin's Student Motivation Scale. Grit is measured with 6 items from the 12-item grit effort scale (Duckworth et al. 2007). Conscientiousness is measured using the NEO-domain 10-item IPIP scale. Data are collected during the first course period among first-year students of the 2018–2019 cohort. Persistence and anxiety are measured during the fifth week of the course period. Grit and Big Five conscientiousness are measured three weeks later.