Peers Affect Personality Development*

Xiaoyue Shan
(National University of Singapore)

Ulf Zölitz
(University of Zurich)

June 2024

Abstract

Do people around us influence our personality? We investigate this question through a field experiment where we randomly assign university students to study groups. We find personality spillovers along three dimensions: students become more conscientious when assigned to conscientious peers, more open-minded when assigned to open-minded peers, and more competitive when assigned to competitive peers. We find no effects for peers' extraversion, agreeableness, or neuroticism. Our findings are consistent with students' adopting peer traits that are predictive of academic achievement. Our paper provides novel evidence on spillovers in noncognitive skills and establishes that socialization with peers affects personality development.

Keywords: personality, spillovers, field experiment, peer effects
JEL classification: I21, I24, J24

* This experiment received IRB approval from the University of Zurich (OEC IRB #2018-021) and was registered at the American Economic Association's registry for randomized controlled trials before the start of the experiment in 2018 (AEARCTR-0003255). In a previous version of this paper (IZA Discussion Paper #15257) we estimated personality spillovers using data from the first three waves of the experiment in 2018/19, 2019/20 and 2020/21. This version uses additional data from replications of the experiment in 2021/22 to 2023/24 and confirms previous results. We received helpful comments from Jan Bietenbeck, Alexandra de Gendre, Bart Golsteyn, Jan Feld, Edwin Leuven, Corinne Low, Nicolás Salamanca, and participants of the CESifo Area Conference on Economics of Education, SOLE 2022, AFE 2022, and the World ESA Meeting 2022, and seminar participants at the University of Zurich and the University of Pennsylvania. We thank Anna Valyogos, Matthew Bonci, Timo Haller, and Francesco Serra for outstanding research assistance. Xiaoyue Shan: National University of Singapore (email: x.shan@nus.edu.sg). Ulf Zölitz: University of Zurich, IZA, CESifo and CEPR (email: ulf.zoelitz@econ.uzh.ch).

1. Introduction

Personality predicts many important outcomes, including education, income, job satisfaction, health, risky behaviors, successful relations, and divorce (Roberts et al. 2007; Borghans et al 2008; Humphries and Kosse 2017). Personality is also a key element of human capital that the labor market increasingly values (Deming 2017; Edin et al. 2022). Given the importance of personality for individuals and society, it is surprising how little we know about causal determinants of personality.

In this paper, we study how peers shape personality. The omnipresence of peers makes it easy to imagine that they influence who we are. This idea is captured by group socialization theory stating that our personality is formed through efforts of fitting into a group and competing with others (Harris 1995). Although peers are promising and seemingly obvious candidates for explaining personality development, causal evidence on their influence is absent. The large literature on peer effects that is devoted to studying social spillovers has never directly investigated this question.

To study the impact of peers on personality development, we conduct a field experiment with 1,229 undergraduate students who we randomly assign to small study groups of four. In these groups, students solve problem sets, prepare tutorial sessions, discuss lectures, and meet for social events. These social interactions take place during the first year at university, a formative period in which students adjust to a new environment, make new friends, and form new habits. The students in our sample are 18–22 years old, an age period where personality still displays substantial malleability (Robins et al. 2001; Caspi and Roberts 2001; Borghuis et al. 2017). We measure students' personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism), with the commonly used Big Five taxonomy. We also measure students' competitiveness, which has recently emerged as an important predictor of education and labor market outcomes (Buser, Niederle, and Oosterbeek 2021). We measure

these six traits at the start of the course before students were assigned to their study groups (baseline) and at the end of the course, just before their final exams (endline). We then estimate how the personality of randomly assigned peers measured at the baseline affects student personality at the endline.

Our results show that students become more similar to their peers along several, but not all, personality dimensions. Being randomly assigned to peers who are one standard deviation (SD) more conscientious raises a student's own conscientiousness by 0.046 to 0.056 SD, depending on the specification. Being assigned to peers who are one SD more competitive makes students 0.067 to 0.078 SD more competitive. We also see that being assigned to peers who are one SD more open to new experiences raises a student's own openness by about 0.063 SD. By contrast, peer extraversion, agreeableness, or neuroticism do not significantly affect a student's own level of the trait.

Are the personality spillovers we document driven by the personality of peers or other characteristics correlated with peer personality? It is hard to make this distinction because peer personality cannot be independently randomized from other peer characteristics. From a policy perspective, this distinction is less important. Knowing that exposure to conscientious peers increases students' conscientiousness is policy relevant, regardless of what drives these effects. In practice, we cannot assign students to more-conscientious peers without changing peer gender, achievement, and other unobserved peer characteristics correlated with conscientiousness. However, to be able to better place our findings in the academic literature, it is important to know whether peer personality is merely a proxy for other peer characteristics that have been shown to predict students' outcomes. We thus test whether controlling for peer

gender, achievement, and a large set of other peer characteristics affects our results. It does not. Having peers with different personalities generates distinct social spillovers.[1]

Our results raise the question of how persistent peer-induced personality changes are. We conduct follow-up surveys and measure personality traits one to four years after the end of the experiment. We find that the spillover effects for conscientiousness and competitiveness remain visible and statistically significant up to four years after the initial peer group assignment. The spillover for openness, however, appears to fade out. The persistent impact of peer conscientiousness and competitiveness suggests that spillovers in these traits go beyond short-term behavioral changes and represent longer-lasting trait changes.

Having established that peers affect personality development, we next investigate whether the three personality traits generating spillovers affect "hard" academic outcomes of other students. We find that exposure to conscientious and competitive peers improves student performance. A one SD increase in peer competitiveness and conscientiousness raises the obtained exam score by 0.083 and 0.046 SD. In comparison, a one SD increase in peers' math ability raises exam performance by 0.044 SD. These findings suggest that having peers with productive noncognitive skills can be as valuable as having high-achieving peers.[2]

Why do peers affect only some personality traits but leave others unaffected? We first test whether peer personality affects the frequency of peer-to-peer interaction. We find that students interact more with conscientious peers. A one SD increase in peers' conscientiousness raises the overall frequency of interaction by about 0.11 SD. By contrast, we find no evidence that peers' competitiveness or openness affects the intensity of academic or social interaction.

[1] We also find no evidence that peers' math ability affects any of the six personality traits, suggesting that peers' cognitive skills do not influence noncognitive skill development.

[2] We also find that having conscientious peers appears to lower students' self-reported anxiety at the end of the semester. A one SD increase in peer conscientiousness lowers the occurrence of anxiety by 3.3 percentage points. This suggests that the effect of peer personality may go beyond personality and academic performance. We also note that personality changes may represent one mechanism for why we observe achievement spillover effects in this and other settings.

These results suggest that enhanced peer-to-peer interaction can be one underlying mechanism of the spillover in conscientiousness, but it cannot explain spillovers for other traits. Next, we test how personality traits relate to academic achievement and find that personality spillovers are only visible for traits that predict educational success. This is consistent with the idea that people engage in self-regulated, effortful personality change (Hennecke et al. 2014; Stieger et al. 2021) and that students only adopt peer traits that are productive in the university context.[3]

The large literature on peer effects has studied how peers' gender, race, or achievement affect performance and educational choices.[4] Only a few recent papers have explored peer personality as an input in the education production function. These studies show that peer personality affects students' performance. Shure (2021) shows that having more conscientious peers raises math and language performance in high school. Hancock and Hill (2021) show that teammate conscientiousness raises team performance in university study groups. Golsteyn, Non, and Zölitz (2021) show that exposure to peers who are more persistent raises university performance.[5] Only one other peer effects paper looks at an outcome related to personality. Using Project STAR data, Bietenbeck (2021) finds that having more motivated peers, while increasing reading test scores, has no significant impact on own motivation. Bietenbeck (2021) studies these effects in the primary school classroom. In contrast, we study peer effects in small university peer groups using six validated personality measures.

[3] We formalize this idea with a simple theoretical framework in Appendix B. In our framework, students adopt traits when the academic returns exceed the costs. The costs of personality change depend on peers who act as role models or apply social pressure.

[4] For example, Hoxby (2000) shows that having more female peers raises both boys' and girls' test scores. Carrell, Sacerdote, and West (2013) highlight that low-achieving students perform worse when medium-achieving peers are replaced with high-achieving peers. Consistent with these results, Booij, Leuven, and Oosterbeek (2017) show that low- and medium-achieving students benefit from tracking of the same type of students. Figlio (2007) shows that boys with female-sounding names have more behavioral problems and a negative impact on their peers' test scores. Carrell, Hoekstra, and Kuka (2018) show that having disruptive peers reduces earnings by 3–4%. Sacerdote (2014) provides an excellent review of the existing literature on peer effects.

[5] In contrast to Golsteyn, Non, and Zölitz (2021), we explicitly test how peer personality affects own personality changes over time and provide supplementary evidence on how peer personality affects exam attendance, performance, and student mental health. While Golsteyn, Non, and Zölitz (2021) exploit institutional random assignment to teaching sections, we conduct a multi-year experiment that explicitly randomizes students into small study groups.

Our work relates to several studies that also conduct experiments to study peer effects (Booij, Leuven, and Oosterbeek 2017; Carrell, Sacerdote, and West 2013; Duflo, Dupas, and Kremer 2011; Oosterbeek and Van Ewijk 2014). While these studies provide important insights into the nature of peer effects, they focus on performance and do not consider personality as an input or output.

This paper underlines the malleability of personality in adolescence and therefore also relates to a series of recent papers showing that targeted interventions can impact socioemotional skills (Alan, Boneva, and Ertac 2019; Abeler, Falk, and Kosse 2021; Alan et al., 2021; Kosse et al. 2020; Sorrenti et al. 2020; Roberts et al. 2017). Our paper contributes to this literature by highlighting that natural exposure to peers can have lasting effects on noncognitive skills. This finding suggests that interventions improving noncognitive skills may have a multiplier effect through peer-to-peer spillovers. List, Momeni, and Zenou (2020) provide evidence along these lines and show that an early childhood intervention affects both treated and untreated children's noncognitive skills. Given these findings and our results, we note that policymakers and researchers might be underestimating the social returns of programs that boost noncognitive skills.

Our paper also has implications for group composition, neighborhood sorting, and peer choice in general. Prior studies have recognized peer effects as underlying mechanisms of the impacts of schools and neighborhoods on socioeconomic outcomes (Jackson 2010; Deming et al. 2014; Chetty, Hendren, and Katz 2016; Deutscher 2020). Peers are also an important factor in families' school and neighborhood choices (Nechyba 2006; Barseghyan, Clark, and Coate 2019). Compared to cognitive ability, demographics, and socioeconomic status, noncognitive skills have been largely neglected. We highlight the value of peers who are equipped with better noncognitive skills and establish a novel fact: peers influence students' personality development.

2. What is Personality and How Much Does It Change?

The American Psychological Association defines personality as "individual differences in characteristic patterns of thinking, feeling and behaving." Heckman and Kautz (2014, page 346) define personality traits as "personal attributes not thought to be measured by IQ tests or achievement tests." Both of these definitions are broad and include socioemotional skills, soft skills, and what economists refer to as "noncognitive" skills. Differentiating personality traits from behaviors, habits, or feelings is not empirically possible (Borghans et al., 2011). Because traits can only be inferred from their consequences such as observable behavior, there can be no direct measurement of an abstract trait. In this paper, we adopt the pragmatic definition by Pervin (1994), who defines personality traits as *observable* patterns of habits and behaviors that make a person unique. Because personality is always measured through questions about behavior, lasting changes in these behaviors represent changes in personality by definition.

What do we know about changes in personality? While there is a misconception among some economists that personality is fixed, a large literature in psychology has shown that personality is both malleable and reasonably stable over time. Roberts et al. (2001) shows that within individuals, personality traits measured at age 18 and 26 show correlations ranging from 0.43 to 0.67. Roberts et al. (2001), Roberts et al. (2006), and Borghuis et al. (2017) show that personality traits remain malleable over the life cycle—conscientiousness, for example, remains fairly malleable until the late twenties. Within economics, a series of recent intervention studies has provided evidence on the malleability of personality during childhood. Kosse et al. (2020) show that children participating in the German "Balu und Du" (Balu and you) mentoring program become more prosocial. Abeler, Falk, and Kosse (2021) show that these children also become more honest, and Boneva et al. (2021) show they become more competitive. In related work, Alan and Ertac (2018) show that children receiving a classroom-based intervention become more patient, and Alan, Boneva, and Ertac (2019) show that grit

can be fostered through interventions. Alan et al. (2021) show that an intervention in perspective-taking increases prosocial behavior. Cappelen et al. (2020) show that early childhood education affects children's social preferences for fairness and the importance children place on efficiency relative to fairness. Sorrenti et al. (2020) show that a socioemotional skills intervention persistently reduces children's impulsiveness and disruptiveness. In related work, Fryer, Levitt, and List (2015) show that incentivized parental engagement can improve children's cognitive and noncognitive scores.

We know substantially less about factors that affect personality in adolescence and adulthood. Dahmann and Anger (2018) study a German schooling reform and find that a one-year reduction in years of schooling increases openness and decreases emotional stability. Only two studies have investigated how schools and teachers affect outcomes related to personality. Jackson (2018) and Jackson et al. (2020) show that both schools and teachers shape socio-emotional development and noncognitive outcomes like absences, suspensions, and on-time grade progression. A more recent work by Joensen et al. (2022) shows that targeted interventions in high school can affect both the level of skill and the technology of skill formation. These results suggest that adolescence can be a formative period for socio-emotional skills.

Recent evidence further suggests that people can change their own personality. Stieger et al. (2021) show that access to a virtual personality coach can help people to persistently change their personality in the desired direction. Experimental participants who had access to such a coach became more conscientious, less neurotic, and more extraverted. Hennecke et al. (2014) propose a framework for self-directed personality development and three necessary conditions to change personality. Based on this framework, people can change their personality if they: (1) feel such a change is desirable or necessary, (2) consider the change to be feasible, and (3) make a habit of the initial changes.

3. Experimental Design

3.1. The Experiment

Our experiment took place in an introductory economics course that is mandatory for all students enrolled in a major or minor in economics, business, or informatics. We conducted the same experiment in the Fall semester for five academic years, from 2018/19 to 2023/24. Students who participated in our experiment were mostly in their first semester and aged 18 to 22 years—a formative period for personality changes (Robins et al. 2001; Caspi and Roberts 2001; Borghuis et al. 2017).

The structure of the course was identical for each of the five student cohorts. Students attended two lectures and one tutorial session per week and completed weekly ungraded problem sets. To pass the course, students needed to pass their final exam, which was the only determinant of their grade. The stakes for this exam were high. If a student failed the exam, it could only be retaken once. Failing the exam for a second time meant the student would have to change majors or drop out of the university.

Figure 1 shows the timeline of our experiment. Before the start of the term, students received an invitation to complete a baseline survey. This survey contained measures of students' personality as well as other demographic and background questions. In the first lecture, we informed students about the possibility of signing up for study groups. While signing up had no direct effect on students' grades, we suggested that they might enjoy studying with other students. Overall, 42% of all students in the course signed up for study groups. We think these students, who are willing to interact with new peers, are arguably a more relevant sample to examine the effect of peer environment on personality development (relative to students who prefer to study by themselves or interact with people out of the social environment being examined). While voluntary participation does not affect our internal validity, it might affect the generalizability of our results. Because we also surveyed students not registered for

study groups, we are able to compare the baseline characteristics by the group registration status. As Table A1 in the Appendix shows, students who signed up were slightly more open to new experiences, more extroverted, and more agreeable. These differences are statistically significant at the 5% level but small in the magnitude: Differences are equivalent to 1–2% of the mean personality level of students who did not register for the course. Along all other dimensions (competitiveness, conscientiousness, neuroticism, high school grades, and gender), participating students were not significantly different from other students (p-value > 0.05). In Section 7 we provide a thorough discussion on the external validity of our results following the framework of List (2020). Taken together, our evidence suggests that our findings may apply to student populations more broadly.

We assigned students to study groups as follows. In each year, we grouped students into three study programs depending on whether they were enrolled in: 1) a business or economics major, 2) an informatics major, or 3) any other major with a business or economics minor.[6] Students who are enrolled in the same broad program typically follow the same curriculum. Within the program, we randomly assigned each student to a study group consisting of four students. Our study group assignment is therefore stratified at the program-by-cohort level. For the COVID-19 cohort in the 2020/21 academic year, the randomization is additionally stratified by the last digit of the student's matriculation number (ID), because the university assigned students to online or in-person lectures based on the last digit of their ID. Since the last digit is a random number by itself, our group assignment remains inherently random for the 2020/21 cohort. We will control for fixed effects for randomization strata (which are smaller for the 2020/21 cohort than other cohorts) in our analysis; see Section 4 for details.

[6] Of them, 65.3% were majoring in business or economics, 15.7% were majoring in informatics, and the remaining 19% were majoring in another subject but were enrolled in a business or economics minor.

We informed students about their study group via email. This email included the names, email addresses, and mobile phone numbers of their study group peers, invited them to create a WhatsApp group, and suggested that they schedule their first group meeting.[7] To foster social interactions, we also offered each group a $20 voucher for drinks at the local university bar. Once students were assigned to groups, participation in group activities was voluntary. Although study groups were designed for studying economics, group members could also study together for other common courses or organize social activities. At the end of the semester, before the final exam, students received the endline survey. This survey elicited students' endline personality traits as well as information about academic and social interactions with their peers.

3.2. Data

We measure students' personality in the baseline survey and in the endline survey. Table 1 provides an overview of the included personality traits, the items used to measure them, and their answer scales.

We measure the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism) with the 15-item Big Five Inventory (BFI-S), which consists of three items per trait (Gerlitz and Schupp 2005). This inventory is a short version of the original 44-item Big Five Inventory (John et al. 1991) and has been validated and used in different settings (Specht et al. 2011; Lang et al. 2011; Hahn et al. 2012). Students rate each item on a 7-point scale ranging from 1 = "does not apply to me at all" to 7 = "applies to me perfectly." Following Gerlitz and Schupp (2005), we measure each personality trait as the average of students' ratings on the three designated items. To measure students' competitiveness, we use the one-item measure proposed and validated by Buser, Niederle, and

[7] According to our pre-experiment survey, more than 90% of students use WhatsApp to communicate with friends, which makes it a convenient tool to facilitate peer group interaction.

Oosterbeek (2021): "*In general, how competitive do you consider yourself?*" Students answer this question on a 10-point scale ranging from 1 = "not competitive at all" to 10 = "very competitive." Appendix Tables A11 and A12 provide an overview of the survey questions.

Our analyses focus on 442 study groups for which we observe all group members' baseline personality traits.[8] Among the 1,776 students in these groups, 1,229 students (69%) completed the endline survey and make up our final estimation sample. Appendix Table A2, Panel B, shows that peer personality at the baseline does not significantly affect the response rate of the endline survey.

Table 2, Panel B, shows descriptive statistics for our estimation sample. Of the 1,229 students, 45.4% are female, approximately 10% are taking the course for the second time, and 87% attended a high school with German as the instruction language.[9] Students also reported their high school final grades in math and the first language using the typical Swiss grading scale, from point 1 to point 6 (higher values indicate better performance).

Panel C shows the summary statistics of personality traits at baseline and endline. From the beginning to the end of the first semester, we only see small changes in students' personality: the average student in our sample becomes more open to new experiences, slightly more extraverted, less competitive, less conscientious, less agreeable, and more neurotic. Appendix Figure A1 plots the distribution of the six personality traits at the baseline and endline.

All six personality traits are empirically distinct from each other. Appendix Table A3 shows that the pairwise correlations between personality traits at the baseline never exceed 0.3 in absolute terms. Table A4 provides evidence on the reliability of our Big Five measures. The Cronbach's Alpha coefficients range from 0.46 to 0.83 and largely support the internal consistency of our personality measures. Table A5 provides evidence on the stability of

[8] We do not include study groups with incomplete baseline personality measures because missing peer data can lead to a large bias in peer-effect estimates (Sojourner 2013).
[9] The instruction language in our setting is German. We coded all our survey questions in both German and English so that students could freely choose the language throughout the survey.

personality over time. The within-trait correlations over time range between 0.62 and 0.80, which is reasonably stable and close to what typically is found in other studies over similar time horizons (Robins et al. 2001; Terracciano, McCrae, and Costa 2010).

Table 2, Panel D, shows two indicators of student performance: exam attendance and exam scores. The final grade for the course is entirely based on student performance in the final exam. Of the students, 97.2% attended the final exam, and conditional on taking the exam, the average score obtained was 60 on a scale of 0 to 120. Considering the disruptive and stressful impacts of the COVID-19 pandemic, we also elicited students' mental health in the endline survey from year 2020/21. The anxiety score summarized in Panel D is based on students' answers to two questions from the Patient Health Questionnaire: *"Over the last two weeks, how often have you been bothered by the following problems? (1) Feeling nervous, anxious or on edge; (2) Not being able to stop or control worrying."* Each question is answered on the following scale: 0 ("Not at all"), 1 ("Several days"), 2 ("More than half of the days"), and 3 ("Nearly every day"). Anxiety is diagnosed based on whether the total score of the two answers reaches 3. Strikingly, we find that 45% of students in our sample have anxiety.

Additionally, Panel D also reports summary statistics for the frequency of academic or social activities of students with their assigned peers. Specifically, from the 2019/20 year onward, we ask students to report their interaction frequency with peers in the group on a scale from 0 ("Never") to 5 ("Multiple times per week"). Since each student has three peers, we calculate the average frequency of academic and social activities at the student level.

4. Empirical Strategy

Our aim is to estimate how peer personality measured at the baseline affects students' own personality at the endline. We estimate the following model separately for each personality trait T:

$$T_{ig,t} = \alpha + \beta \bar{T}_{-ig,t-1} + \gamma_1 W_{ig,t-1} + \gamma_2 X_{ig,t-1} + \varepsilon_{igt}, \quad (1)$$

where $T_{ig,t}$ is the trait level of student *i* in study group *g* measured in the endline survey at time *t*. $\bar{T}_{-ig,t-1}$ is the average trait level of students in group *g* (excluding student *i*) measured in the baseline survey at time *t−1*. In our main analysis, we estimate Equation (1) separately for each personality trait. For example, we estimate the effect of having peers who are more open on students' own openness.

The vector $W_{ig,t-1}$ refers to *randomization controls* including student cohort dummies and randomization strata dummies. As explained in Section 3.1, for all regular cohorts, the strata dummies refer to study program dummies: 1) business or economics majors, 2) informatics majors, and 3) the base group: other majors. For the cohort of 2020/21, there are 29 dummies (three times ten minus one) because the randomization is stratified by the study program and the last digit of matriculation numbers.[10] The vector $X_{ig,t-1}$ includes control variables that differ by specification. In all specifications, $X_{ig,t-1}$ is the baseline level of the personality trait in question. For example, we control for students' own openness at the baseline when our dependent variable is students' openness at the endline.

We estimate additional specifications where $X_{ig,t-1}$ includes students' other personality traits measured at the baseline (*other own personality traits*), student's gender, age, fixed effects for business-economics majors, high school math and language grades, high school study hours per week, and an indicator for whether German was the instruction language in high school (*other own characteristics*), as well as peer averages of these variables (*other peer personality traits* and *other peer characteristics*). To facilitate the interpretation of our estimates, we standardize each personality trait to have a mean of zero and a standard deviation

[10] Considering the uniqueness of the COVID-19 cohort in terms of the randomization process and learning conditions, we test the robustness of our results by excluding this cohort from our sample. Results in Appendix Table A6 show that the estimates of personality spillovers remain very similar.

of one across the estimation sample. The peer averages reported in our regression tables are based on those standardized personality measures.[11]

The key identifying assumption for our approach is that the randomization of students into study groups was successful. To confirm that this is the case, we test how peer personality measured at the baseline relates to students' own personality measured at the *baseline*, that is, before peer composition could have affected it. We implement this test by estimating Equation (1) but use as the dependent variable students' baseline personality instead of their endline personality. Besides randomization controls, these regressions include a leave-out-mean of each personality trait at the program-by-cohort level to account for the mechanical relationship between own and peer characteristics (see Guryan, Kroft, and Notowidigdo 2009).

Table 3 shows that peer personality at the baseline does not significantly predict students' own personality at baseline. All six estimates of interest are small in magnitude, and none are statistically significant at the 5% level.[12] Table A7 in the Appendix provides an alternative balancing check in which we test whether study group dummies jointly predict students' baseline characteristics. This is not the case. Both tests confirm that our randomization of students into study groups was successful.

5. Results

5.1. Personality Spillovers

Figure 2 shows that students become more similar to their randomly assigned peers for three out of six personality traits. Being assigned to peers who are one SD more competitive increases students' own competitiveness by 0.067 SD (*p*-value = 0.022). Being assigned to

[11] Appendix Figure A2 visualizes where our identifying variation comes from. The difference between groups with above- and below-median levels of peer conscientiousness is 0.9, which is equivalent to 1.6 individual level standard deviations.

[12] Peer conscientiousness negatively predicts own conscientiousness, an effect marginally significant (*p*-value= 0.098). In practice this should make it harder for us to find positive spillover effects for conscientiousness. Note that throughout our analysis we always control for the own baseline personality trait.

peers who are one SD more open to new experiences increases students' own openness by 0.060 SD (p-value = 0.019). Similarly, being assigned to peers who are one SD more conscientious increases students' own conscientiousness by 0.046 SD (p-value = 0.019). We find no evidence that being assigned to more extraverted, agreeable, or neurotic peers affects students' own personality in these dimensions.

How large are these effects in their original scales? In terms of the 10-point scale used to measure competitiveness, being assigned to peers who are, on average, one point more competitive increases students' own competitiveness by 0.12 points. In terms of the 7-point scale used to measure the Big Five personality traits, we see that being assigned to three peers who are one point more open increases own openness by 0.11 points, and being assigned to peers who are one point more conscientious increases own conscientiousness by 0.09 points. These effects are modest but economically meaningful.

We assess the robustness of these results by gradually adding student- and peer-level control variables. Column (1) of Table 4 shows our baseline estimates from Figure 2 for reference. Column (2) shows estimates from regressions in which we include all own personality variables in the model, and in column (3), we additionally include all student background variables. As expected under random assignment, including these variables does not affect our point estimates in any meaningful way.

Are these effects driven by peer personality? Or, do they show effects of characteristics correlated with peer personality? These questions are difficult to answer because we cannot randomly assign peer personality independent of other peer attributes. For policymakers, understanding what drives our observed effects is also less important. They can, for example, increase students' competitiveness by assigning them to more competitive peers *regardless of what drives these effects*.

To be able to better place our results in the literature, it is nevertheless important to know whether our peer personality spillovers merely capture effects associated with other peer variables, for example, peer gender and peer achievement, which are the two peer characteristics most frequently studied in the literature. We therefore test how our point estimates are affected by including other peer variables. In column (4) of Table 4, we include all peer personality variables in the model, and in column (5), we include peer gender, high school math and language grades, and whether German was the high school instruction language. Our results remain very similar after including these controls. This suggests that our measures of peer personality capture distinct components of students' personality traits.[13] Column (5) shows that peer conscientiousness increases own conscientiousness (p-value = 0.023), peer competitiveness increases own competitiveness (p-value = 0.024), and peer openness increases own openness (p-value = 0.020). Taken together, our estimates in Table 4 show that the impact of peer personality remains remarkably robust and is not sensitive to the inclusion of student- or peer-level control variables.

While our conclusions remain the same, it is not clear whether controlling for other peer characteristics is a good robustness check. The estimates from specifications that include these controls do not deliver policy-relevant parameters. In practice, we cannot assign students to more-conscientious peers without changing peer gender, achievement, or other correlated unobserved peer characteristics.

Our results raise the question whether institutions should systematically measure student personality and if these measures are useful for admission decisions and groups assignment. While it would be valuable to measure personality at a large scale, we can expect that students will provide strategic answers when it becomes known that their responses will

[13] In column (5), the estimated effect of peers' high school math grades on all own personality traits are statistically insignificant. This finding suggests that peers' cognitive skills have limited impact on the development of noncognitive skills.

be used for admissions or peer group assignments. For example, a prospective student may try to appear particularly hardworking when answering items about conscientiousness hoping to be admitted or to receive hardworking peers. Until new forms of personality measurement like the implicit personality test,[14] which are very hard to game, solve these issues, it is not clear whether one may want to base education policies on these survey-based personality measures.

Full Spillover Matrix: Although this paper focuses on within-trait personality spillovers, for completeness, we also test whether spillovers exist across traits. Appendix Table A9 provides a "full spillover matrix" and shows estimates for how each peer trait affects the six own traits. We find that personality spillovers are highly concentrated on the diagonal line, where a given own trait is regressed on the same peer trait. However, six out of the other 30 coefficients reach the 5% significance level, which could represent chance findings.

Since we are estimating 36 effects simultaneously, the results in the full spillover matrix are likely susceptible to multiple hypothesis testing. We therefore apply multiple testing corrections to the 36 estimates.[15] After the correction, the within-trait spillovers in competitiveness, openness, and conscientiousness remain significant at the 5% or 1% level. While effects on the diagonal line do survive the multiple testing correction, significant effects elsewhere do seem to largely disappear once the correction is applied. Only one of the six significant effects that are not on the diagonal survives the correction.

Measurement Error: One might be concerned that measurement error in the peer personality variables drives our results. If measurement error is random, this should not be a concern. Feld

[14] An implicit personality test is designed to measure personality traits that participants may be strategically unwilling to reveal in conventional tests (Grumm and von Collani, 2007).
[15] We use the step-down procedure of Romano and Wolf (2005, 2016) to control for the familywise error rate simultaneously for the 36 hypotheses. We use bootstrapping to resample group observations (resampling clusters) for 10,000 times – standard errors are also clustered at the group level in all original regressions. See Section 7.1 for further discussions of multiple hypothesis testing.

and Zölitz (2017) show that random measurement error will attenuate peer effect estimates in settings with random assignment. If peer personality measures true personality with random error, we would therefore expect our results to be a lower bound of the true effect.

5.2 Convergence of Personality

Our analyses reveal spillovers in competitiveness, openness, and conscientiousness: when exposed to peers with higher (lower) levels of a trait, own levels of the trait at the endline increase (decrease). An alternative interpretation of these positive spillover effects is that students' own personality converges to peer personality. Put differently, students become more similar to their peers along these dimensions of personality.

To better illustrate the patterns of personality convergence, we calculate the distance between own personality T_0^o and peer personality T_0^p at baseline for each trait ($d_0 = T_0^o - T_0^p$), and similarly, the distance between own personality at endline T_1^o and peers' baseline personality T_0^p ($d_1 = T_1^o - T_0^p$). If students do converge to peers in terms of a trait, the endline distance should move closer to zero compared to the baseline distance. Except in the case of overshooting, the two distances should also have the same sign. In Figure 3, we plot d_1 against d_0 separately for each trait. The results show that along the dimensions of competitiveness, openness, and conscientiousness, students do move closer to T_0^p—the reference points set by their peers at the baseline. For these three traits, the endline distance is significantly smaller than the baseline distance ($d_1 < d_0$). By contrast, for the other three traits, the endline and baseline distances are not significantly different. If anything, we find suggestive evidence of divergence along the dimension of neuroticism ($d_1 > d_0$).

5.3 Heterogeneity in Spillover Effects

In this section, we examine whether the spillovers are driven by certain types of students. We use binned scatter plots to show how own personality varies with peer personality in Figure 4. Figure 4 shows that spillover effects in competitiveness, openness, and conscientiousness are fairly linear. Furthermore, we separately estimate the impact of peer personality for students with above- and below-median values of a given trait at the baseline. Results in Appendix Figure A3 show that the impact does not significantly vary with students' own personality. Given the lack of evidence on heterogeneity, we think the linear-in-means model is a good approximation to describe spillovers in our setting.

5.4 Persistence of Personality Changes

We have shown that peers affected students' personality at the time of the endline survey, which was three months after the study group assignment. What happens after the course is over? Personality peer effects may fade out and students may revert to their old selves once they are no longer exposed to their study group peers, or peer effects may persist if students have formed new habits or continue to interact with their peers after the end of the course.

To provide evidence on the persistence of peer personality effects, we conducted two follow-up surveys in the summer of 2021 and 2022 to measure the personality of students from the previous cohorts (2018/19 to 2021/22). Depending on their cohort, students were surveyed one to four years after the end of the experiment. The average response rate was 46% for the two follow-up surveys. Reassuringly, peer traits generating spillovers are unrelated to the probability of responding to this survey.[16]

[16] Table A10 in the Appendix tests whether baseline peer personality predicts follow-up survey participation. We find no evidence that any of the six peer personality traits significantly predicts the response rate.

Figure 5 shows how peer personality at baseline relates to own personality at the baseline, endline, and follow-up. For ease of comparison, we use raw scores of personality outcomes (instead of standardized measures) on a scale of 1 to 7 or 1 to 10 for both the independent and dependent variables. The baseline estimates in Figure 5 restate that peers' and own traits are not significantly correlated, confirming that peers are randomly assigned to study groups. We then re-estimate our main results from column (3) of Table 4. The point estimates are different because we use raw personality measures.

Finally, we examine how personality traits measured years later respond to peer personality measured at baseline and plot the estimates in the last column of Figure 5. Overall, we find suggestive evidence of persistent personality spillovers. For competitiveness and conscientiousness, the follow-up spillovers are less precisely estimated but have comparable or even somewhat larger effect sizes compared to the endline spillovers. Up to four years after being assigned to a study group, students who had peers who were one point more competitive are 0.2 points more competitive (compared to 0.12 points at the endline). Similarly, students who had peers who were one point more conscientious are 0.13 points more conscientious (compared to 0.08 points at the endline). For openness, we see lower point estimates at the follow-up than at the endline (0.034 vs. 0.114 points), but the 95% confidence interval includes the endline effect size. Taken together, these results suggest that spillovers in conscientiousness and competitiveness are more persistent, while spillovers in openness are relatively short lived.

Although the follow-up effects are less precisely estimated and therefore more suggestive, they give us confidence that our main results are not driven by short-term behavioral changes. Our results suggest that peers can induce lasting changes in behavior that remain visible in personality measures.

5.5 Impact of Peer Personality on Performance and Mental Health

Does peer personality also affect students' academic performance and well-being? Even though peers' academic ability does not affect own personality development, peer personality may still affect own academic performance. Such effects could be a result of peer-induced changes in students' own personality, direct effects of peer personality on performance, or both. Peer personality may also affect students' mental health due to the quality of social interaction itself or indirectly through academic and personality spillovers. In this section, we analyze whether the three traits generating personality spillovers also affect course performance and anxiety levels at the end of the semester.

Overall, we see some evidence that peer personality impacts performance. Panels A and B of Figure 6 show that some peer traits affect the likelihood that students attend the final exam and the final exam score. Students exposed to peers who were one SD more open to new experiences become 0.8 percentage points less likely to take the final exam (p-value = 0.036). Students with peers who were one SD more competitive perform 0.082 SD better on the final exam (p-value = 0.001). Similarly, peers who were one SD more conscientious improve exam performance by 0.050 SD (p-value = 0.051). These results provide one possible explanation for why the spillovers in competitiveness and conscientiousness are more long-lasting while the spillover in openness fades away over time. The positive effects on both personality and performance may reinforce each other and lead to longer-term behavioral changes.

Panel C of Figure 6 shows that exposure to more-conscientious peers also improves students' mental health. Specifically, students with peers who were one SD more conscientious are 3.3 percentage points less likely to experience anxiety (p-value = 0.045). Conscientious peers tend to be more reliable, organized, and responsible. Having peers with these traits may therefore be beneficial for work collaboration in study groups, leading to lower levels of stress and better academic performance.

As a benchmark, we also estimate the impact of peers' pre-university math achievement on performance and mental health. We find that exposure to high-achieving peers has a marginally significantly positive impact on exam performance but does not significantly impact the probability of exam attendance or anxiety levels. When comparing the effects of peer academic ability and peer personality, we find that peer competitiveness and conscientiousness create even slightly stronger effects on performance (although the differences are not statistically different). Meanwhile, peer conscientiousness has a strong positive effect on mental health, which is statistically significantly different from the effect of peer academic ability.

Taken together, our findings suggest that exposure to peers with productive noncognitive skills is potentially more important than exposure to high-achieving peers. Our finding that peer personality affects own personality and performance, while peer achievement only marginally affects performance, suggests that the peer-to-peer spillovers of cognitive and noncognitive skills on each other are potentially asymmetric. This echoes the skill formation model of Cunha and Heckman (2008), which highlights that noncognitive skills promote the formation of cognitive skills, but cognitive skills mostly do not promote the formation of noncognitive skills.

6. Mechanisms

Peers impact students' personality development. We find spillovers for conscientiousness and competitiveness that appear to persist and spillovers for openness that seem more short-lived. These results raise the question of why the spillovers are concentrated on certain traits. Given the large literature on socialization and peer effects, we find it unsurprising that peers *can* affect

personality development, especially during a formative period in early adulthood.[17] The finding that seems puzzling to us is that peers induce personality change for some traits, but leave other traits unaffected. To understand the sources of trait-specific spillovers, we explore two possible mechanisms: the intensity of peer-to-peer interaction and the motivation for personality change.

6.1 Peer-to-Peer Interaction

We first examine whether peers with different personality traits induce different degrees of social interaction. If students interact more with peers who are more competitive, open-minded, and conscientious, the personality spillovers may simply be due to more intense exposure to peers with these traits. To test this, we use data on the frequency of interaction measured in the endline survey. More specifically, starting from the year 2019/20, students reported how frequently they had interacted with their study mates to pursue academic or social activities. Answers were recorded on a scale from 0 ("Never") to 5 ("Multiple times per week").

Table 5 shows how different peer personality traits affect the frequency of academic and social interactions. The dependent variables are the standardized frequency of academic interaction (column 1), social interaction (column 2) and the overall frequency of interaction (column 3), which combines the two former measurements. Perhaps unsurprisingly, having more extroverted peers increases the frequency of joint social activities. We also find that conscientious peers significantly increase the frequency of joint learning activities and the overall frequency of interaction. By contrast, peers' competitiveness or openness does not

[17] Broadly classified, the existing literature provides three possible channels for why peers *can* affect personality development. The role model or *social learning channel* captures the idea that students learn from peers' characteristics, behaviors, and outcomes, and adapt their own personalities (Bandura and Walters 1963; Bandura and McClelland 1977; Moretti 2011, Bursztyn et al. 2014). The *social comparison channel* states that students use peers as reference points to compare themselves to and become more similar to their peers' personalities (Suls et al. 2002; Rayo and Becker 2007; Chen et al. 2010). The *group socialization channel* states that group members' personalities become more like each other due to within-group assimilation or the pressure of group norms (Harris 1995; Rubin et al. 2006; Reitz et al. 2014).

significantly affect peer-to-peer interactions. If anything, competitive peers seem to lower the frequency of joint learning and social activities. Altogether, these results suggest that increased peer-to-peer interaction can be one underlying mechanism of the spillovers of conscientiousness, but it is unlikely to drive spillovers of other personality traits.

6.2 Self-Motivated Personality Change

One reason spillovers are trait specific could come from students having different incentives to change specific personality traits. McCrae and Costa (1994) propose that personality change is costly, and Hennecke et al. (2014) argue that people may only engage in changes that they consider desirable and feasible. Translating these ideas to our context implies that we would expect personality spillovers to be concentrated on traits that are valuable to students in our academic setting.[18] Therefore, we analyze the academic returns to different personality traits to understand the *relevance* of each trait. More specifically, we test how different personality traits correlate with two indicators of academic achievement that we observe in our setting: the high school grade and the final grade for the economics course in university.

Figure 7 shows that at baseline, conscientiousness and competitiveness are the traits that significantly predict high school performance. A one SD increase in own conscientiousness (competitiveness) predicts a 0.078 SD (0.184 SD) increase in high school grades. At endline, openness and conscientiousness become the most relevant traits. Students who are one SD less open to new experiences at baseline obtain grades that are 0.172 SD higher in the course. Those who are one SD more conscientious at baseline obtain grades that are 0.102 SD higher. Baseline

[18] We formally illustrate this intuition behind this cost-benefit analysis of personality change with a simple theoretical framework in Appendix B. In this framework, a personality trait may or may not affect academic performance, and changing a trait is costly. The presence of peers affects the cost of personality change due to social learning or pressure. Our key prediction is that students become more like their peers for traits that affect academic performance. For traits that are irrelevant for academic performance, students lack the incentive to change.

extraversion and neuroticism also appear to predict the final economics performance, but the magnitude is much smaller compared openness and conscientiousness.[19] Taken together, these results highlight that personality spillover effects are only visible for traits that predict performance. These patterns are consistent with the idea that students engage in self-motivated personality change for traits that are *academically relevant*.

7. Robustness and Discussions

7.1 Robustness of Results

Extreme Values of Personality: One potential concern for the interpretation of our results is ceiling effects. For example, students with the highest trait levels can only move downward, which may appear like they are becoming more like their peers. To mitigate the concern, we test whether our results are driven by students displaying extremely low or high levels of personality traits at baseline. In Appendix Figure A4, we show spillover effects for the full sample and the sample of students whose baseline personality traits do not display extreme values. We find that the point estimates and confidence intervals remain very similar after dropping the extreme values. These results suggest that ceiling effects do not affect our estimates.

Multiple Hypothesis Testing: Even though we conduct the same analyses in our previous working paper, concerns about multiple hypothesis testing remain. Figure 2 and Figure 6 highlight that we look at multiple outcomes and multiple peer characteristics, including the spillovers of six personality traits and how three peer traits affect educational and mental health

[19] When examining the correlations between *endline* personality traits and the economics grade, we find that competitiveness also significantly predicts the final performance. A one SD increase in competitiveness predicts a 0.11 SD increase in the course grade (p-value < 0.001). Similarly, openness and conscientiousness at endline also significantly predict the economics grade.

outcomes. Therefore, we conduct multiple testing corrections *simultaneously* for these 15 estimates, even though this paper focuses on the six estimates of personality spillovers.

We use two methods to adjust the *p*-values of main estimation results. First, we use the approach of Romano and Wolf (2005, 2016) with resampling procedures to control the family-wise error rate (FWER). Specifically, when implementing the Romano–Wolf correction, we resample clusters of groups within cohort and study program strata 10,000 times and keep clustering standard errors at the group level for the original regression models.[20] Second, we control the false discovery rate (FDR) and calculate the sharpened *q*-values (Anderson 2008; Benjamini, Krieger, and Yekutieli 2006). Table 6 shows the main results of this paper, including the estimated coefficients, the original *p*-values [in square brackets], the Romano–Wolf *p*-values {in curly brackets}, and the FDR adjusted *q*-values (in parentheses). We find that the within-trait spillovers of competitiveness, openness, and conscientiousness remain statistically significant at the 5% level when using either of the two correction methods. The impact of peer openness on exam attendance and the impact of peer conscientiousness on exam performance remain marginally significant. In addition, the impact of peer competitiveness on performance and the effect of peer conscientiousness on anxiety remain statistically significant at the 5% level. These results mitigate concerns over multiple hypothesis testing.

7.2 Pre-registration and Replication of the Experiment

Pre-Registration: Our initial pre-registration (AEARCTR-000325) was filed before the start of the experiment in 2018/19 and was not very specific about the statistical tests we would perform and did not have a formal pre-analysis plan. In our IZA working paper (Shan and

[20] The correction results are very similar when modifying the resampling procedure—for example, when resampling student-level observations instead of group-level observations. Results are also very similar when we use the baseline sample (N = 1,776) with all four group members to resample group clusters.

Zölitz 2022), we analyzed personality spillovers using the first three waves of data collection (2018/19 to 2020/21). To rule out that the initial results represented chance findings, we replicated our experiment.

Replication: After releasing our working paper that reported results based on three experimental cohorts, we replicated the experiment in three additional years (2021/22 to 2023/24). Throughout this paper we combine data from all six cohorts and follow the empirical specification as in our previous working paper. Figure 8 shows our main results (the six estimates of personality spillover) separately for the first three waves (the initial sample), the last three waves (the replication sample), and all six waves (the pooled sample). The results appear very similar both qualitatively and quantitatively between the initial sample and the replication sample. While the point estimate for conscientiousness is somewhat smaller in the replication, we overall view these results as strong evidence that our results successfully replicate and do not represent chance findings.

7.3 External Validity

This paper uses a field experiment to examine peer-to-peer personality spillovers among first-year university students. Are our results externally valid beyond our experimental setting? Following the guidelines provided by List (2020), we discuss the external validity of our results from the following perspectives: sample selection, attrition, naturalness, and scaling.

Selection: Our experimental subjects are first-semester university students who are interested in group work and social interactions. As discussed above, students who signed up for study groups in the class were similar to other students with respect to most baseline characteristics including gender, achievement, competitiveness, agreeableness, neuroticism, etc. We find that students who registered were slightly less conscientious and more open-minded. This suggests

that our analysis sample is fairly representative of all students in the setting, except that the nonparticipants seem to have a lower demand for interactions with peers. It is also reassuring that our results are robust to controlling for all baseline personality traits and do not significantly vary with the levels of baseline personality.

Are students at the Swiss university we study representative of young adults in other contexts and other countries? To answer this question, we compare personality traits of our sample to respondents to the 2019 German Socio-Economic Panel (G-SOEP), aged 18 to 22. The key advantage of this comparison is that the G-SOEP used the exact same items and answer scale to measure the Big Five personality traits in a representative German sample (SOEP 2020). Figure 9 plots the means and 95% confidence intervals of each personality trait separately for the whole sample of G-SOEP respondents (N = 25,979), the respondents aged 18 to 22 (N = 1,920), all students in our setting who reported their personality at baseline (N = 5,232), and our final analysis sample of students who signed up for study groups and satisfied the inclusion criteria mentioned in Section 3.2 (N = 1,229). We find that students in our sample are very similar to the German representative sample of young adults (and even the whole sample) along all the dimensions of personality—both in terms of level and variability.

Attrition: Our analysis focuses on students who reported their personality both at the baseline and endline. Over the period of a semester, the attrition rate was about 30%. However, we find no evidence on selective attrition: the endline response rate does not vary with experimental conditions. Appendix Table A2 shows that peer personality does not affect attrition. To study educational outcomes, we use administrative data and therefore face no attrition issues related to these outcomes. Finally, as Table 5 shows, we find no evidence for selective compliance with the experiment interventions: students' likelihood and intensity of interactions with assigned peers do not vary with peer personality.

Naturalness: Our experiment took place in a first-semester university class, where studying in groups and interacting with new peers were normal. Random (study) group assignment is also a common practice employed in schools and universities (Webb 1982); for recent examples in higher education see Feld and Zölitz (2017) and Hancock and Hill (2021). The baseline and endline surveys were framed as general course surveys and sent to all students in the class—not only the experimental sample with assigned study groups. Therefore, subjects were not aware that their reported personality scores would be used for analyses on peer effects. The exogenous variations we examine in this paper represent typical natural variations in group composition that would occur in many other settings. In other words, we use a naturally designed field experiment without artificial interventions.

Scaling: Given the high naturalness of our experiment, we think our analyses can be easily replicated in other settings. We study variations in group composition that often occur naturally. Therefore, to examine the scaling feature of our results, the relevant question is not whether the program can be implemented in the general population, but rather whether our estimated spillovers can be detected in other settings. As highlighted in our model, personality spillovers are concentrated in traits that are relevant for individuals' goals in a given context. Thus, we think that personality spillovers might be to some degree context-specific. For example, in art studies, where open-mindedness is more valuable, we may find spillovers in openness. In theater and acting studies where extraversion is more productive, we may detect spillovers in extraversion. We also expect that it will be more difficult to detect personality spillovers in later life stages when personality is less malleable, or in weak or fragmented social networks with less peer-to-peer interaction.

Taken together, our findings are based on a naturally designed field experiment without obvious selection or attrition issues. Our sample is also representative of young adults in a general population regarding personality traits. Therefore, we think that our results have high external validity and can apply to similar youth and adult populations that learn through social interactions. We expect to find different patterns of personality spillovers in other contexts where returns to personality and costs of personality change differ from our setting. Future studies could explore the transmission of personality and noncognitive skills in other contexts to help us better understand personality development under the influence of peers.

7.4 Study Group Treatment Effect

Do students in a study group perform better than students without a study group? Our study is not designed to answer this question because in our setting the decision to sign up for a study group is clearly endogenous. Figure A5 nevertheless provides some descriptive evidence on this question. Figure A5 shows student performance for those who signed up for a study group compared to those that did not. When using the raw course grade, we find that students with a study group perform on average 0.16 SD better. After controlling for a wide range of student characteristics such as student gender, high school grades, and baseline personality measures, this performance gap is reduced to 0.07 SD.

8. Conclusion

Attending university is formative for students. They socialize with their peers and adapt to their new environment. This paper represents the first systematic study on how the peer environment shapes students' personality. To identify the causal impact of peer personality, we conduct a field experiment that randomizes undergraduate students into study groups that have frequent social interactions.

We find that the peers students meet at the beginning of their studies have a lasting impact on their personality development. These spillovers are trait specific. Students increase their competitiveness, openness, and conscientiousness if their study group consists of peers who score higher on these traits. Extraversion, agreeableness, and neuroticism appear unaffected by peer personality. Consistent with previous studies, we provide suggestive evidence that peer personality also influences student performance.

We further explore why peers affect some traits but not others. We find that the trait-specific spillovers are not generally driven by more-intensive interactions with peers carrying certain personality traits. Instead, by examining how different traits correlate with academic performance, we find that the traits creating spillovers are more "academically relevant" to students in the setting. Taken together, these results suggest that students adopt peer traits that are valuable to them and help them succeed in the academic environment.

Our paper establishes that personality is malleable and shaped through socialization with peers. It provides the first causal link between the peer environment and personality development. The existence of these personality spillovers is important for policymakers and practitioners in education who assign students to classes or groups, employers who allocate workers to teams, families who choose schools or neighborhoods, and students making choices of peers or friends. Our results suggest that exposure to peers with productive personality traits can be as important as exposure to high-performing peers. Given the growing returns to noncognitive skills in education and the labor market (Deming 2017; Edin et al. 2022), it is becoming increasingly important to understand the causal determinants of personality change and to consider personality as an explicit policy target. Our results also raise the question of whether economic and social preferences are similarly malleable. Future work could study whether the peer environment affects the development of risk, time, social or other preferences.

References

Abeler, J., Falk, A., & Kosse, F. (2021). Malleability of preferences for honesty. *CESifo Working Paper 9033.*

Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, *136*(4), 2147-2194.

Alan, S., Boneva, T., & Ertac, S. (2019). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, *134*(3), 1121-1162.

Alan, S., & Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy*, *126*(5), 1865-1911.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association, 103*(484), 1481-1495.

Bandura, A., & Walters, R.H. (1963). *Social Learning and Personality Development*. Holt Rinehart and Winston: New York.

Bandura, A., & McClelland, D. C. (1977). *Social Learning Theory* (Vol. 1). Prentice Hall, Englewood Cliffs.

Barseghyan, L., Clark, D., & Coate, S. (2019). Peer preferences, school competition, and the effects of public school choice. *American Economic Journal: Economic Policy*, *11*(4), 124-58.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491-507.

Bietenbeck, J. (2021) Peer Motivation, and Educational Success. *Working Paper*.

Boneva, T., Buser, T., Falk, A., & Kosse, F. (2021). The origins of gender differences in competitiveness and earnings expectations: Causal evidence from a mentoring intervention. *IZA Discussion Paper No. 14800.*

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *The Review of Economic Studies*, *84*(2), 547-578.

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

Borghans, L., Golsteyn, B. H., Heckman, J., & Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences*, *51*(3), 315-320.

Borghuis, J., Denissen, J. J. A., Oberski, D., Sijtsma, K., Meeus, W. H. J., Branje, S., Koot, H. M., & Bleidorn, W. (2017). Big Five personality stability, change, and codevelopment across adolescence and early adulthood. *Journal of Personality and Social Psychology, 113*(4), 641-657.

Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, *82*(4), 1273-1301.

Buser, T., Niederle, M., & Oosterbeek, H. (2021). Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures. *NBER Working Paper No. w28916.*

Cappelen, A., List, J., Samek, A., & Tungodden, B. (2020). The effect of early-childhood education on social preferences. *Journal of Political Economy, 128*(7), 2739-2758.

Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry, 12*(2), 49-66.

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.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review, 108*(11), 3377–3415.

Chen, Y., Harper, F. M., Konstan, J., & Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on MovieLens. *American Economic Review, 100*(4), 1358-1398.

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review*, *106*(4), 855-902.

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, *43*(4), 738-782.

Deutscher, N. (2020). Place, peers, and the teenage years: Long-run neighborhood effects in Australia. *American Economic Journal: Applied Economics, 12*(2), 220-49.

Dahmann, S. C., & Anger, S. (2018). Cross-fertilizing gains or crowding out? Schooling intensity and noncognitive skills. *HCEO Working Paper* 2018-065.

Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review, 104*(3), 991-1013.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593-1640.

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*(5), 1739-74.

Edin, P. A., Fredriksson, P., Nybom, M., & Öckert, B. (2022). The rising return to noncognitive skill. *American Economic Journal: Applied Economics*, *14*(2), 78-100.

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 Finance and Policy*, *2*(4), 376-394.

Fryer, R. G., Levitt, S. D., & List, J. A. (2015). Parental incentives and early childhood achievement: A field experiment in Chicago heights. *NBER Working Paper No. 21477*.

Gerlitz, J. Y., & Schupp, J. (2005). Zur Erhebung der Big-Five-basierten persoenlichkeitsmerkmale im SOEP. *DIW Research Notes*, *4*, 2005.

Golsteyn, B. H., Non, A., & Zölitz, U. (2021). The impact of peer personality on academic achievement. *Journal of Political Economy*, *129*(4), 1052-1099

Grumm, M., & von Collani, G. (2007). Measuring Big-Five personality dimensions with the implicit association test–Implicit personality traits or self-esteem?. *Personality and Individual Differences*, *43*(8), 2205-2217

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

Hancock, S., & Hill, A. (2021) The effect of teammate personality on team production. *Working Paper*.

Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality – Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, *46*(3), 355-359.

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review*, *102*(3), 458-489.

Heckman, J. J., & Kautz, T. (2014). Fostering and measuring skills: Interventions that improve character and cognition. In Heckman, J. J., Humphries, J. E., & Kautz, T. (Eds.). *The*

myth of achievement tests: The GED and the role of character in American life. (pp. 341-430). University of Chicago Press.

Hennecke, M., Bleidorn, W., Denissen, J. J., & Wood, D. (2014). A three–part framework for self–regulated personality development across adulthood. *European Journal of Personality*, *28*(3), 289-299.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Paper No. 7867*.

Humphries, J. E., & Kosse, F. (2017). On the interpretation of non-cognitive skills–What is being measured and why it matters. *Journal of Economic Behavior & Organization*, *136*, 174-185.

Jackson, C. K. (2010). The effects of an incentive-based high-school intervention on college outcomes. *NBER Working Paper No. 15722*.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072-2107.

Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights*, *2*(4), 491-508.

Joensen, J. S., List, J. A., Samek, A., & Uchida, H. (2022). Using a field experiment to understand skill formation in the adolescent years. *Working Paper*.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big five inventory. *Journal of Personality and Social Psychology*.

Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H., & Falk, A. (2020). The formation of prosociality: Causal evidence on the role of social environment. *Journal of Political Economy*, *128*(2), 434-467.

Lang, F. R., John, D., Lüdtke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, *43*(2), 548-567.

List, J. A. (2020). Non est disputandum de generalizability? A glimpse into the external validity trial. *NBER Working Paper No. 27535*.

List, J. A., Momeni, F., & Zenou, Y. (2020). The social side of early human capital formation: Using a field experiment to estimate the causal impact of neighborhoods. *NBER Working Paper No. 28283*.

McCrae, R. R., & Costa Jr, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, *3*(6), 173-175.

Moretti, E. (2011). Social learning and peer effects in consumption: Evidence from movie sales. *The Review of Economic Studies*, *78*(1), 356-393.

Nechyba, T. J. (2006). Income and peer quality sorting in public and private schools. *Handbook of the Economics of Education*, 2, 1327-1368.

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

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, *5*(2), 103-113.

Rayo, L., & Becker, G. S. (2007). Habits, peers, and happiness: An evolutionary perspective. *American Economic Review*, *97*(2), 487-491.

Reitz, A. K., Zimmermann, J., Hutteman, R., Specht, J., & Neyer, F. J. (2014). How peers make a difference: The role of peer groups and peer relationships in personality development. *European Journal of Personality*, *28*(3), 279-288.

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81*(4), 670-683.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1-25.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313-345.

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*(2), 117-141.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality*, *69*(4), 617-640.

Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica, 73*(4), 1237-1282.

Romano, J. P., & Wolf, M. (2016). Efficient computation of adjusted *p*-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, *113*, 38-40.

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 571-645). John Wiley & Sons, Inc.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics, 6*(1), 253-272.

Shan, X., & Zölitz, U. (2022). Peers affect personality development. IZA Discussion Paper No. 15257. Available at https://docs.iza.org/dp15257.pdf

Shure, N. (2021). Non-cognitive peer effects in secondary education. *Labour Economics, 73*, 102074.

SOEP Innovation Sample (SOEP-IS), data from 1998-2019. 2020. DOI: 10.5684/soep.is.2019

Sojourner, A. (2013). Identification of peer effects with missing peer data: Evidence from Project STAR. *The Economic Journal*, *123*(569), 574-605.

Sorrenti, G., Zölitz, U, Eisner, M. & Ribeaud, D. (2020). The causal impact of socio-emotional skills training on educational success. *IZA Discussion Paper No. 13087.*

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, *101*(4), 862-882.

Stieger, M., Flückiger, C., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2021). Changing personality traits with the help of a digital personality change intervention. *Proceedings of the National Academy of Sciences*, *118*(8) e2017548118.

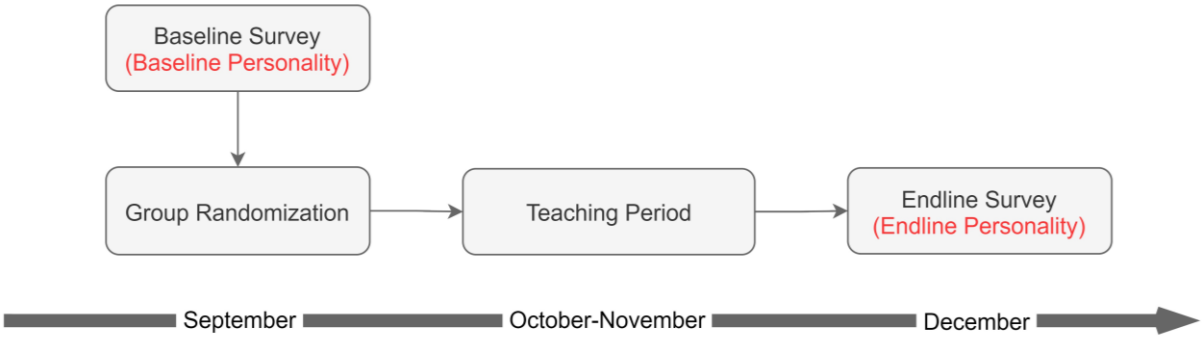
Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, *11*(5), 159-163.

Terracciano, A., McCrae, R. R., & Costa Jr, P. T. (2010). Intra-individual change in personality stability and age. *Journal of Research in Personality*, *44*(1), 31-37.

Webb, N. M. (1982). Student interaction and learning in small groups. *Review of Educational Research, 52*(3), *421-445.*

TABLES AND FIGURES

Figure 1: Timeline of the Experiment



Notes: We conducted the same experiment among six cohorts from 2018/19 to 2023/24. The experiment timeline has remained the same across the years.

Table 1: Measurement of Personality Traits (C-OCEAN)

Personality Trait	Question and Scale
Competitiveness	**In general, how competitive do you consider yourself?** *Scale: 0 (not competitive at all) to 10 (very competitive)*
	I see myself as someone who . . . *Scale: 1 (does not apply to me at all) to 7 (applies to me perfectly)*
Openness	is original, comes up with new ideas values artistic, aesthetic experiences has an active imagination
Conscientiousness	does a thorough job tends to be lazy [reversed] does things effectively and efficiently
Extraversion	is communicative, talkative is outgoing, sociable is reserved [reversed]
Agreeableness	is sometimes somewhat rude to others [reversed] has a forgiving nature is considerate and kind to others
Neuroticism	worries a lot gets nervous easily is relaxed, handles stress well [reversed]

Notes: Students could choose to fill out the survey in German or English. More than 80% of students answered the survey in German. The German version of the 15-item Big Five personality traits was taken from the German Socio-Economic Panel (G-SOEP). The single-item competitiveness scale is based on an early version of Buser, Niederle, and Oosterbeek (2021).

Table 2: Descriptive Statistics

	(1)	(2)	(3)	(4)	(5)
Panel A: Distribution across Cohort & Study Program					
Cohort 2018/19	7.6%	Business & economics majors			66.5%
Cohort 2019/20	14.9%	Banking & finance			16.7%
Cohort 2020/21	22.1%	Business administration			37.3%
Cohort 2021/22	17.3%	Economics			12.5%
Cohort 2022/23	17.5%	Informatics majors			16.2%
Cohort 2023/24	20.6%	Business & economics minors			17%
	N	mean	sd	min	max
Panel B: Background Characteristics					
Female	1,229	0.454	0.498	0	1
Age	1,229	20.19	1.328	18	22
Retaking the Course	1,229	0.102	0.302	0	1
High School Math Grade	1,229	4.590	0.799	1	6
High School Language Grade	1,229	4.749	0.593	1	6
High School in German Language	1,229	0.868	0.338	0	1
High School Study Hours	1,229	9.885	10.160	0	87
Panel C: Student Personality					
Baseline Openness	1,229	4.824	1.114	1.333	7
Endline Openness	1,229	4.881	1.151	1	7
Baseline Conscientiousness	1,229	4.813	0.961	1.333	7
Endline Conscientiousness	1,229	4.671	1.038	1.667	7
Baseline Extraversion	1,229	4.748	1.232	1	7
Endline Extraversion	1,229	4.756	1.278	1	7
Baseline Agreeableness	1,229	5.504	0.860	2.667	7
Endline Agreeableness	1,229	5.326	0.919	2.333	7
Baseline Neuroticism	1,229	4.043	1.325	1	7
Endline Neuroticism	1,229	4.170	1.379	1	7
Baseline Competitiveness	1,229	6.923	1.931	1	10
Endline Competitiveness	1,229	6.650	2.077	1	10
Panel D: Outcomes					
Exam Attendance	1,229	0.972	0.164	0	1
Exam Score	1,195	59.81	20.17	8	118
Anxiety	953	0.451	0.498	0	1
Frequency of Academic Interaction	1,127	1.214	1.266	0	5
Frequency of Social Interaction	1,127	0.283	0.694	0	5

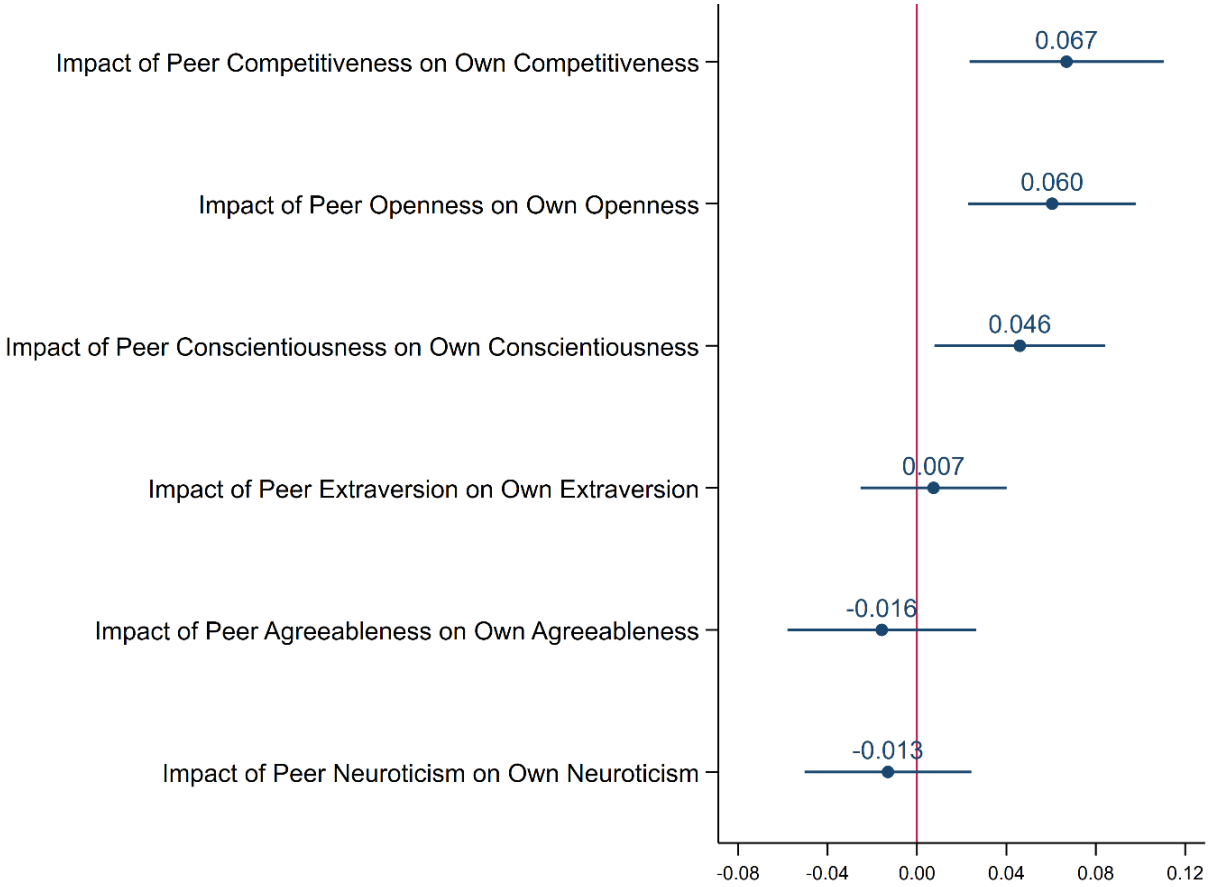
Notes: This table is based on our estimation sample. The standard deviation is denoted by "sd." The sample size is smaller for exam score, anxiety, and interaction frequency in Panel D because the exam score is conditional on taking the final exam, the indicator for anxiety is only available from 2020/21, and the frequency of interaction is only available from 2019/20.

Table 3: Randomization Test

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Own Personality Traits at Baseline					
	Competitiveness	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Peer Competitiveness	-0.004					
	(0.029)					
Peer Openness		-0.030				
		(0.030)				
Peer Conscientiousness			-0.051*			
			(0.031)			
Peer Extraversion				-0.040		
				(0.036)		
Peer Agreeableness					-0.035	
					(0.031)	
Peer Neuroticism						0.048
						(0.032)
Observations	1,229	1,229	1,229	1,229	1,229	1,229
R-squared	0.332	0.317	0.320	0.307	0.315	0.293

Notes: Each column represents one OLS regression. All regressions control for study program-by-cohort fixed effects and the leave-out-mean of each personality trait at the program-by-cohort level. All dependent and independent variables are standardized. Robust standard errors clustered at the study group level are in parentheses. Panel A of Appendix Table A2 shows the balancing results using the full baseline sample (1,776 students in 442 groups). In Appendix Table A8, we show that the six peer personality traits do not jointly predict any of the six own personality traits at baseline. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 2: The Impact of Peer Personality at Baseline on Own Personality at Endline



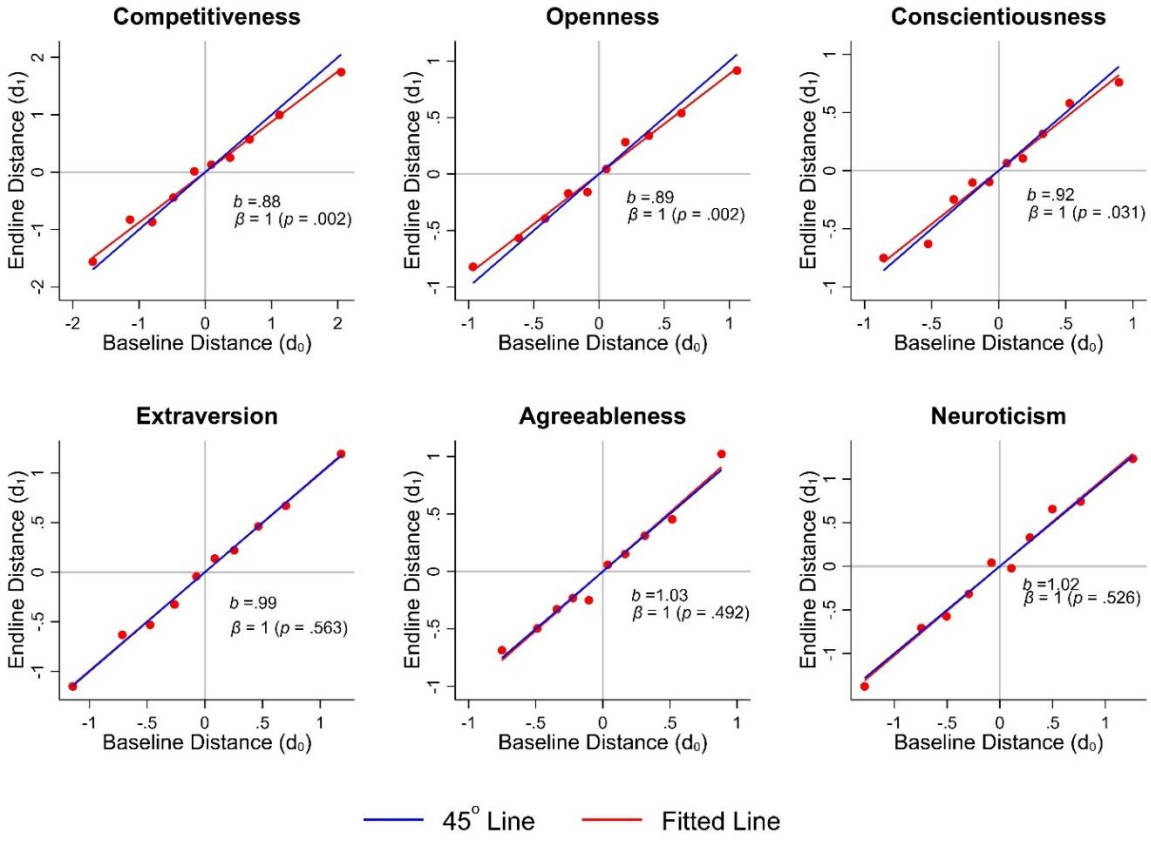
Notes: Each point estimate represents one OLS regression as specified in column (1) of Table 4. For each regression, the outcome variable is a standardized own personality trait at the endline, the independent variable of interest is a standardized peer personality trait at the baseline, and the control variables include randomization controls as defined in the empirical strategy and students' own personality trait at the baseline (i.e., the baseline level of the dependent variable). Standard errors are clustered at the study group level. Error bars indicate 95% confidence intervals.

Table 4: Peer Personality and Own Personality

	(1)	(2)	(3)	(4)	(5)
Panel A: DV = Std. Own Competitiveness at Endline					
Std. Peer Competitiveness	0.067***	0.068***	0.069***	0.078***	0.078***
	(0.022)	(0.022)	(0.022)	(0.024)	(0.024)
	[0.003]	[0.002]	[0.002]	[0.001]	[0.001]
Panel B: DV = Std. Own Openness at Endline					
Std. Peer Openness	0.060***	0.062***	0.061***	0.072***	0.063***
	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)
	[0.002]	[0.001]	[0.002]	[0.000]	[0.002]
Panel C: DV = Std. Own Conscientiousness at Endline					
Std. Peer Conscientiousness	0.046**	0.044**	0.042**	0.054**	0.056**
	(0.019)	(0.020)	(0.019)	(0.021)	(0.023)
	[0.019]	[0.025]	[0.031]	[0.011]	[0.015]
Panel D: DV = Std. Own Extraversion at Endline					
Std. Peer Extraversion	0.007	0.005	0.004	0.009	0.010
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)
	[0.657]	[0.763]	[0.827]	[0.606]	[0.579]
Panel E: DV = Std. Own Agreeableness at Endline					
Std. Peer Agreeableness	-0.016	-0.016	-0.015	-0.008	-0.010
	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)
	[0.463]	[0.467]	[0.493]	[0.714]	[0.645]
Panel F: DV = Std. Own Neuroticism at Endline					
Std. Peer Neuroticism	-0.013	-0.013	-0.012	-0.015	-0.019
	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)
	[0.494]	[0.507]	[0.526]	[0.449]	[0.350]
Observations	1,229	1,229	1,229	1,229	1,229
Control Variables:					
Other own personality traits	N	Y	Y	Y	Y
Other own characteristics	N	N	Y	Y	Y
Other peer personality traits	N	N	N	Y	Y
Other peer characteristics	N	N	N	N	Y

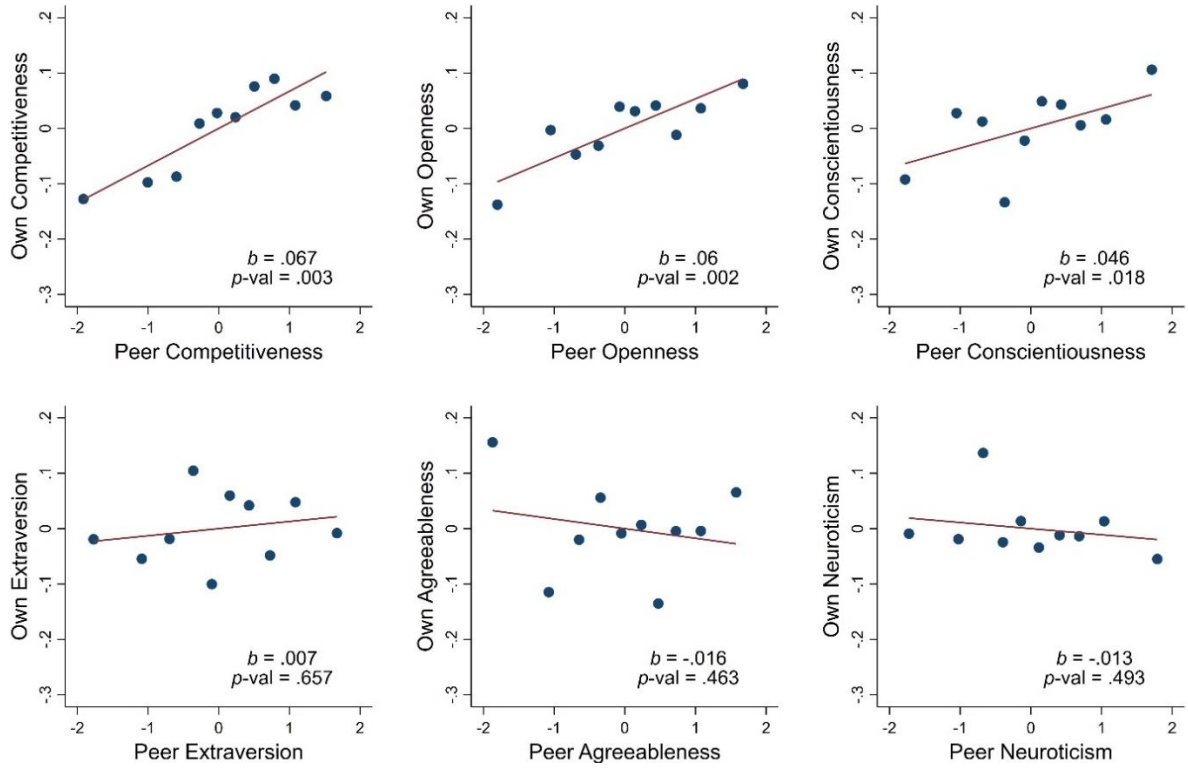
Notes: Each column in each panel shows one coefficient from a separate OLS regression. All regressions include the baseline level of the dependent variable as well as randomization controls as defined in the empirical strategy. *Other own personality traits* include five other personality traits at the baseline, except for the one trait being examined in each panel. *Other own characteristics* include gender, business-economics major fixed effects, high school math and language grades, high school study hours, an indicator for German as the high school instruction language, course-retaking status, and age fixed effects. *Other peer personality traits* are five other peer personality traits at the baseline, except for the one being analyzed. *Other peer characteristics* include peers' gender, high school math and language grades, and whether German was the instruction language in high school. All dependent and independent variables are standardized. The peer averages reported are based on standardized personality measures. Robust standard errors clustered at the study group level are in parentheses. *p*-values are in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 3: Convergence of Personality



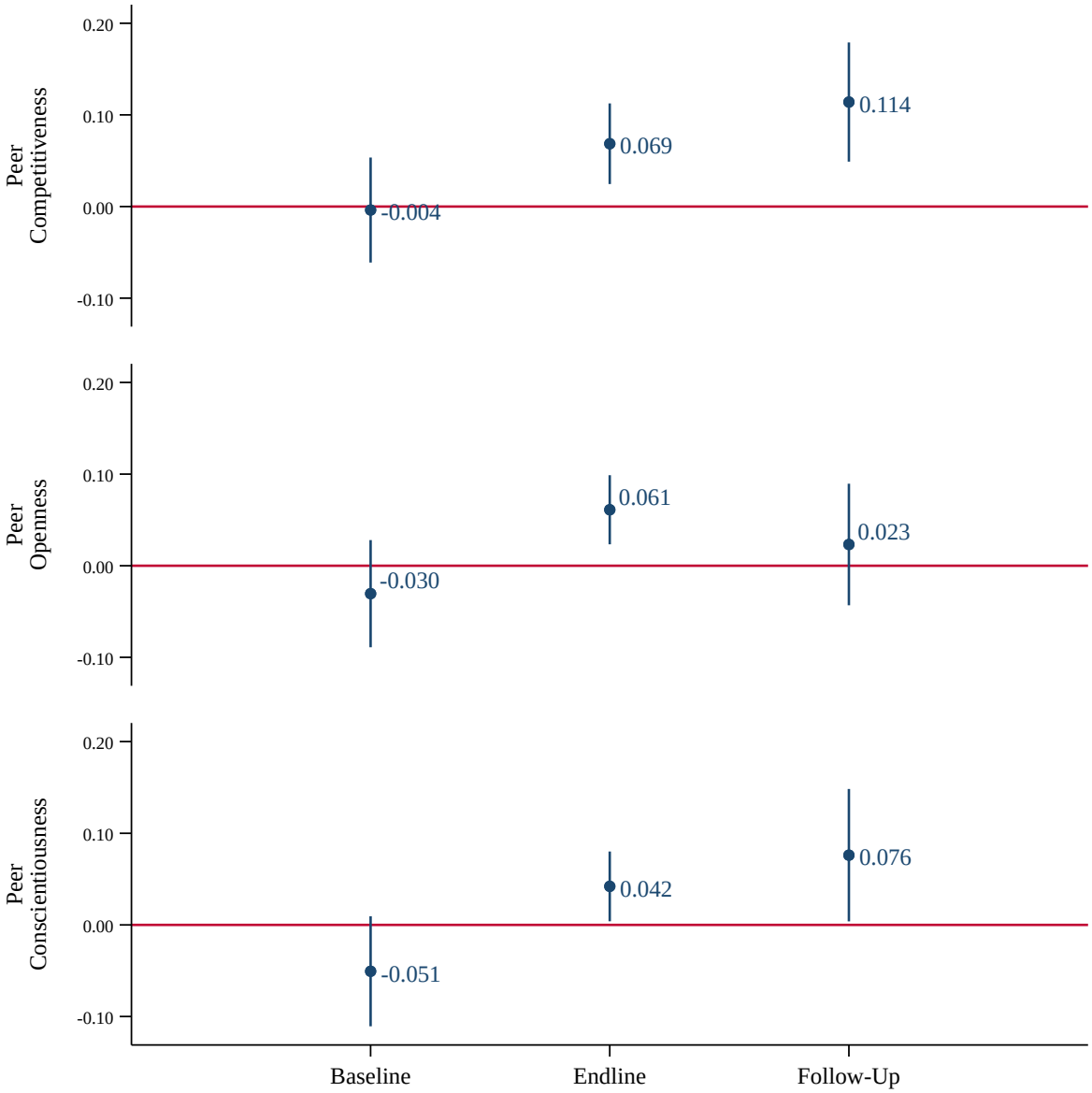
Notes: Binned scatter plots showing the relationship between the endline and baseline distance. The baseline distance (d_0) is the difference between a student's own level of a personality trait and peers' level of this trait at the baseline. The endline distance (d_1) is the difference between a student's own level at the endline and peers' level at the baseline. All baseline and endline distances are the residuals after controlling for individual characteristics observed at the baseline. The 45° line is a reference line representing the case of no convergence to peer personality: $d_1 = d_0$. The graph also plots the fitted linear lines and the estimated slopes (b). We also test whether the estimated slopes are significantly different from one ($H_0: \beta = 1$) and show the *p*-values for the tests.

Figure 4: Non-Linear Spillovers?



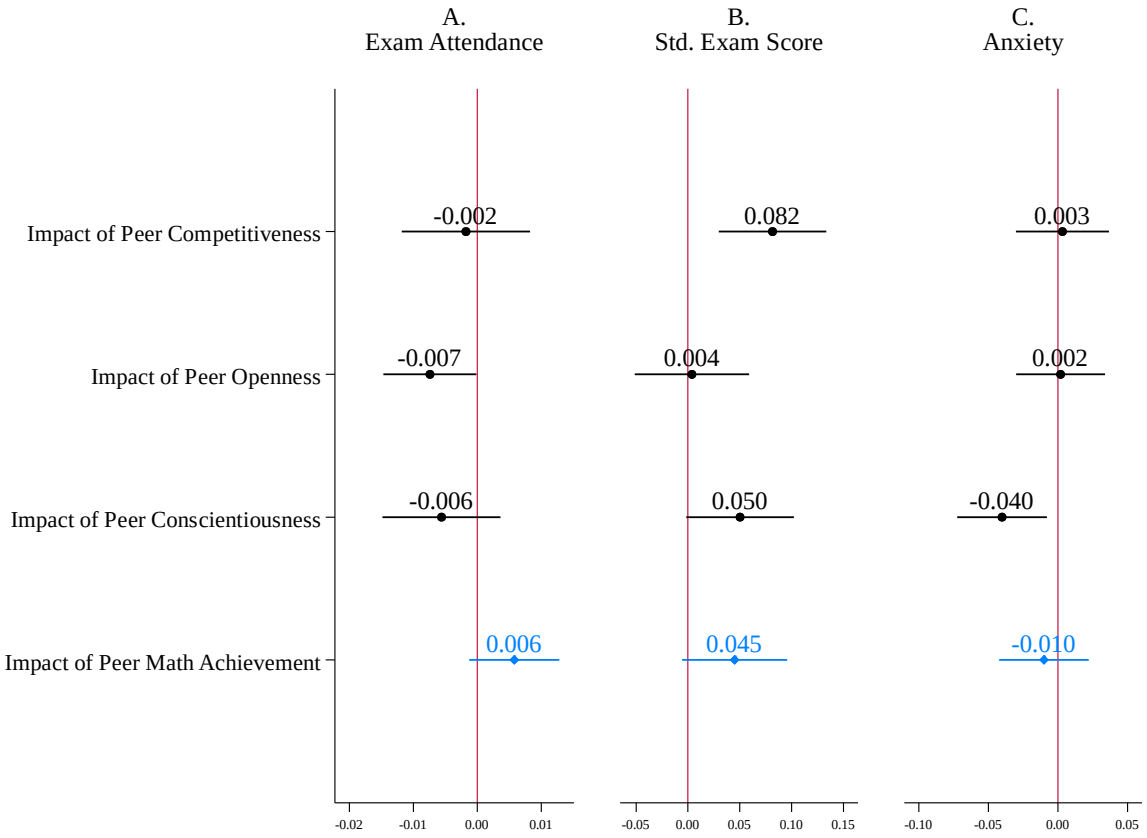
Notes: Binned scatter plots showing the relationship between own personality traits at endline (y-axis) and peers' personality traits at baseline (x-axis). All measurements of own and peer personality traits are standardized. Each plot uses the same specification as in column (3) of Table 4.

Figure 5: Balance, Short- and Longer-Term Effects of Peer Personality



Notes: The figure shows estimated effects of peer personality on students' own personality measured at the baseline, in the endline survey, and in the 2021 and 2022 follow-up surveys—up to four years after the experiment. The estimation sample includes participants from the first four cohorts (2018/19 to 2021/22). Baseline estimates use the same specifications outlined in Table 3; the endline estimates are based on the model reported in column (3) of Table 4. To estimate the impact of peers' baseline personality on student personality in the longer run, we append measures in both follow-up surveys for cohorts from 2018/19 to 2020/21 and use the 2022 follow-up data for the 2021/22 cohort. All specifications include the same control variables as the specification in column (3) of Table 4 and dummies for the respective follow-up year. To make sure that the outcomes are comparable across time, we use raw scores of personality traits at baseline, endline, and follow-up on a scale of 1 to 10 or 1 to 7. Error bars indicate 95% confidence intervals.

Figure 6: Impact of Peer Personality and Achievement on Educational Outcomes



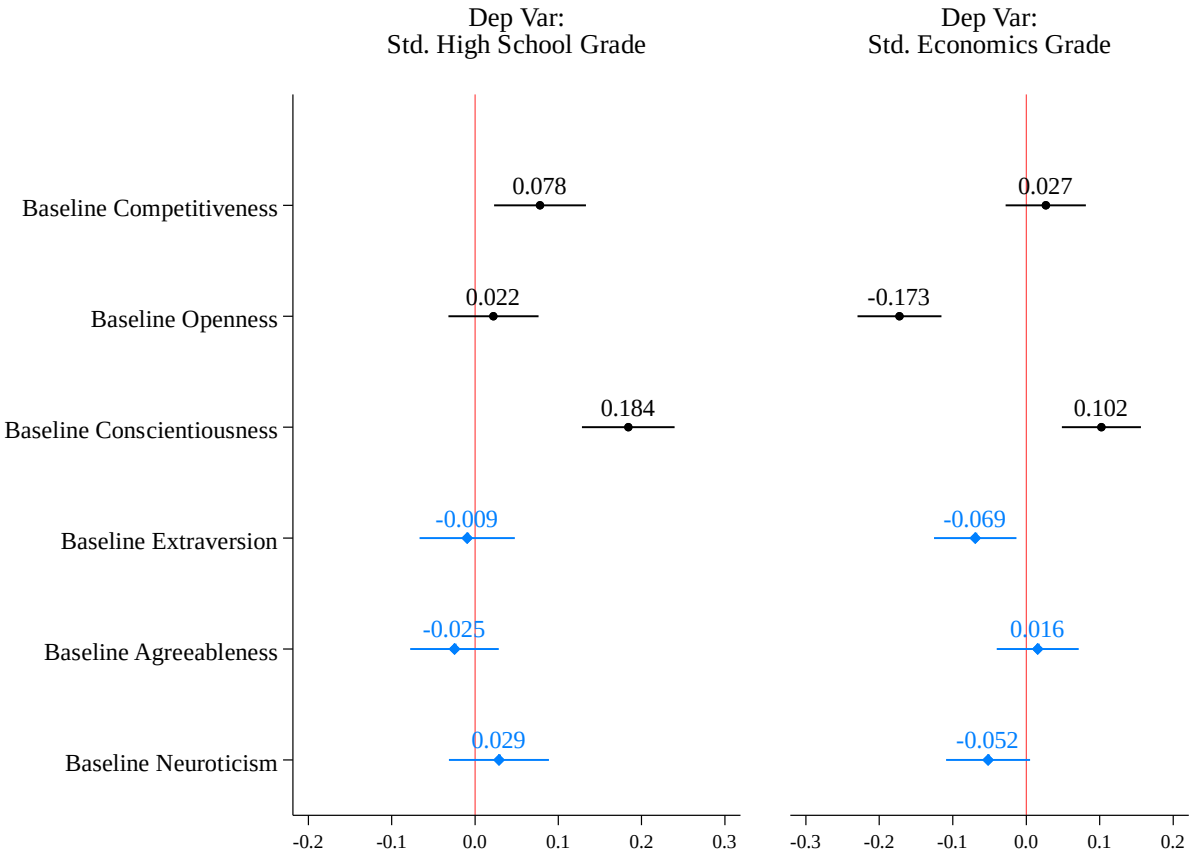
Notes: The figure shows how peer personality affects exam attendance and the obtained exam score for the economics course and reported occurrence of anxiety at the end of the semester. The exam score is standardized with a mean of zero and a standard deviation of one. We also examine how peers' math achievement affects the three outcomes. Each point estimate represents one OLS regression. All specifications include the same controls as column (1) of Table 4. Error bars indicate 95% confidence intervals.

Table 5: Peer Personality and Interaction Frequency

	(1) Std. Frequency of Academic Interaction	(2) Std. Frequency of Social Interaction	(3) Std. Overall Frequency of Interaction
Peer Competitiveness	-0.038	-0.029	-0.040
	(0.049)	(0.039)	(0.044)
Peer Openness	0.009	-0.014	-0.003
	(0.046)	(0.045)	(0.047)
Peer Conscientiousness	0.110**	0.070	0.109**
	(0.048)	(0.044)	(0.047)
Peer Extraversion	-0.020	0.094**	0.044
	(0.047)	(0.042)	(0.044)
Peer Agreeableness	-0.059	-0.015	-0.045
	(0.043)	(0.046)	(0.047)
Peer Neuroticism	0.025	0.002	0.017
	(0.048)	(0.043)	(0.047)
Observations	1,126	1,126	1,126
R-squared	0.127	0.105	0.121

Notes: We test whether peer personality impacts the frequency of academic interaction with peers (studying together for classes) in column (1), the frequency of social interaction (joint extracurricular activities) in column (2), and the overall frequency of interaction with peers in column (3). The peer-to-peer interaction data were collected in the endline survey for cohorts from 2019/20 to 2023/24, so the sample size is smaller. The frequency of academic and social interaction is measured with two survey questions with an answer scale from 0 ("Never") to 5 ("Multiple times per week"). The overall frequency in column (3) is a standardized measure, with a mean of zero and a standard deviation of one, combining the frequency of academic and social interaction. Each column represents one regression, using the same specification as in column (4) of Table 4. All peer traits and interaction frequency variables are standardized among the analysis sample. Standard errors clustered at the group level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 7: Correlation Between Personality and Academic Achievement



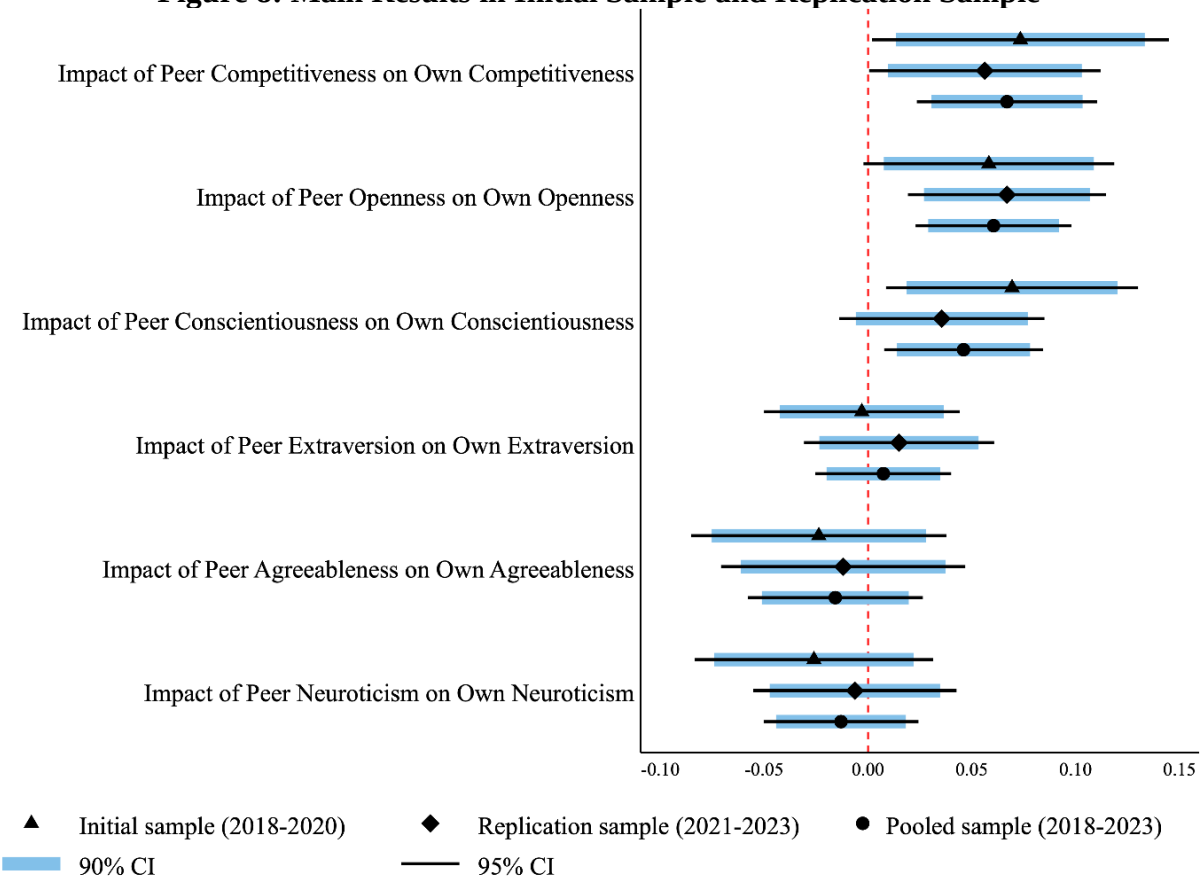
Notes: The figure shows how baseline personality traits predict the high school grade (the average of math and language grades) and the final grade for the economics course where we conducted the experiment. The raw measures of both grades are on a scale of 1 to 6. We use a standardized measure of them (with a mean of zero and a standard deviation of one) in the analysis. Estimates are based on 12 separate regressions controlling for cohort and study program fixed effects. Error bars indicate the 95% confidence intervals based on robust standard errors.

Table 6: Multiple Testing Correction for Main Results

	(1)	(2)	(3)	(4)
	Within-Trait Spillovers	Impact on Performance and Anxiety		
		Exam Attendance	Std. Exam Score	Anxiety Indicator
Peer Competitiveness	0.067	-0.002	0.082	0.003
	[0.003]	[0.726]	[0.002]	[0.851]
	{0.002}	{0.982}	{0.001}	{0.990}
	(0.014)	(0.778)	(0.014)	(0.778)
Peer Openness	0.060	-0.007	0.004	0.002
	[0.002]	[0.045]	[0.892]	[0.908]
	{0.001}	{0.059}	{0.990}	{0.990}
	(0.014)	(0.083)	(0.778)	(0.778)
Peer Conscientiousness	0.046	-0.006	0.050	-0.040
	[0.018]	[0.233]	[0.058]	[0.015]
	{0.017}	{0.579}	{0.075}	{0.013}
	(0.047)	(0.305)	(0.091)	(0.047)
Peer Extraversion	0.007			
	[0.657]			
	{0.982}			
	(0.778)			
Peer Agreeableness	-0.016			
	[0.463]			
	{0.929}			
	(0.653)			
Peer Neuroticism	-0.013			
	[0.493]			
	{0.929}			
	(0.653)			

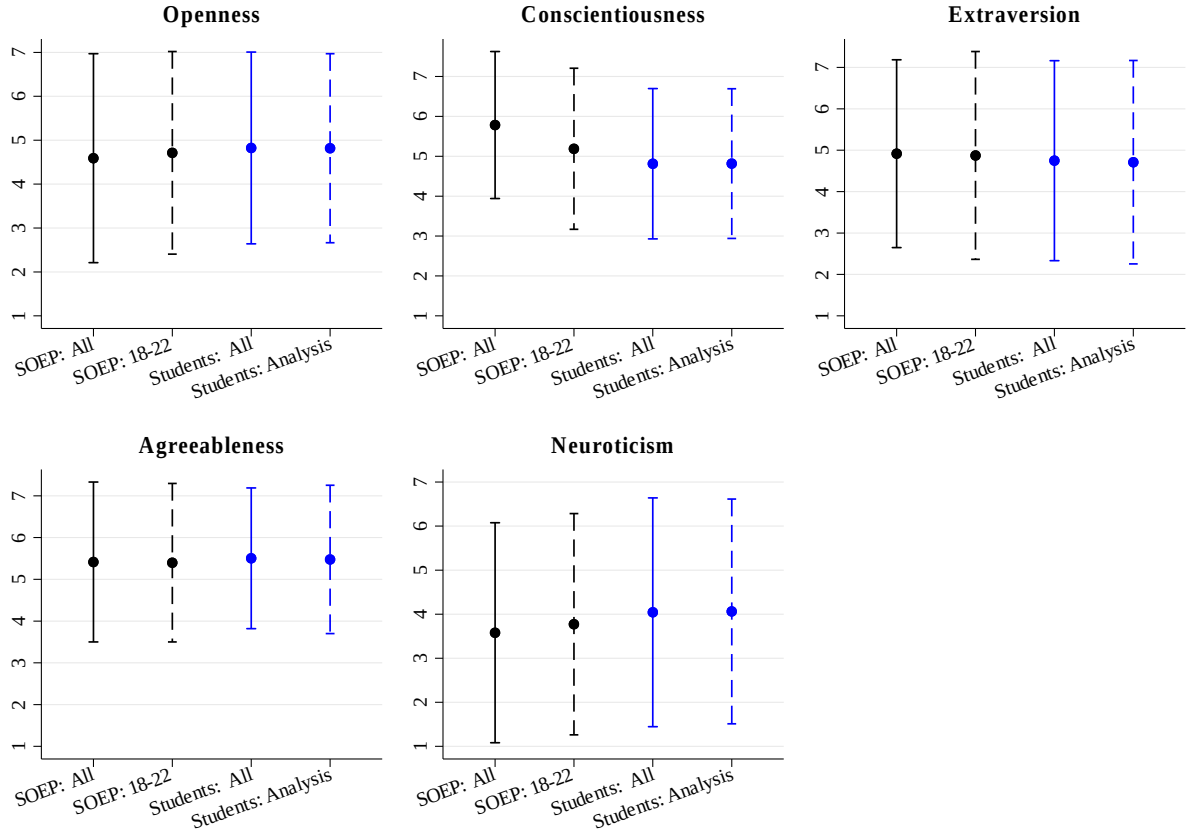
Notes: The table shows the original and corrected *p*-values for the main results of this paper—the original results are presented in column (1) of Table 4 and Figure 6. Each estimate is derived from one OLS regression. We show the estimated coefficients and the original *p*-values based on clustered standard errors in square brackets. Then, we correct the *p*-values by controlling the family-wise error rate (FWER), using the approach of Romano and Wolf (2005, 2016) with 10,000 resampling of group clusters within cohort-by-program strata. When implementing the Roman–Wolf correction, we keep using clustered standard errors for the original models. The corrected *p*-values are in curly braces. For robustness, we also calculate the false discovery rate (FDR) adjusted *q*-values using the method of Anderson (2008) and show the *q*-values in parentheses.

Figure 8: Main Results in Initial Sample and Replication Sample



Notes: The figure shows the estimated within-trait spillover effects using the initial sample (2018/19 to 2020/21) included in the IZA discussion paper No.15257, the replication sample (2021/22 to 2023/24), and the pooled sample (2018/19 to 2023/24). Each estimate is derived from one OLS regression that uses the same specification as in Figure 2. Error bars indicate 90% and 95% confidence intervals (CI).

Figure 9: Personality Traits of Our Student Sample versus SOEP Respondents

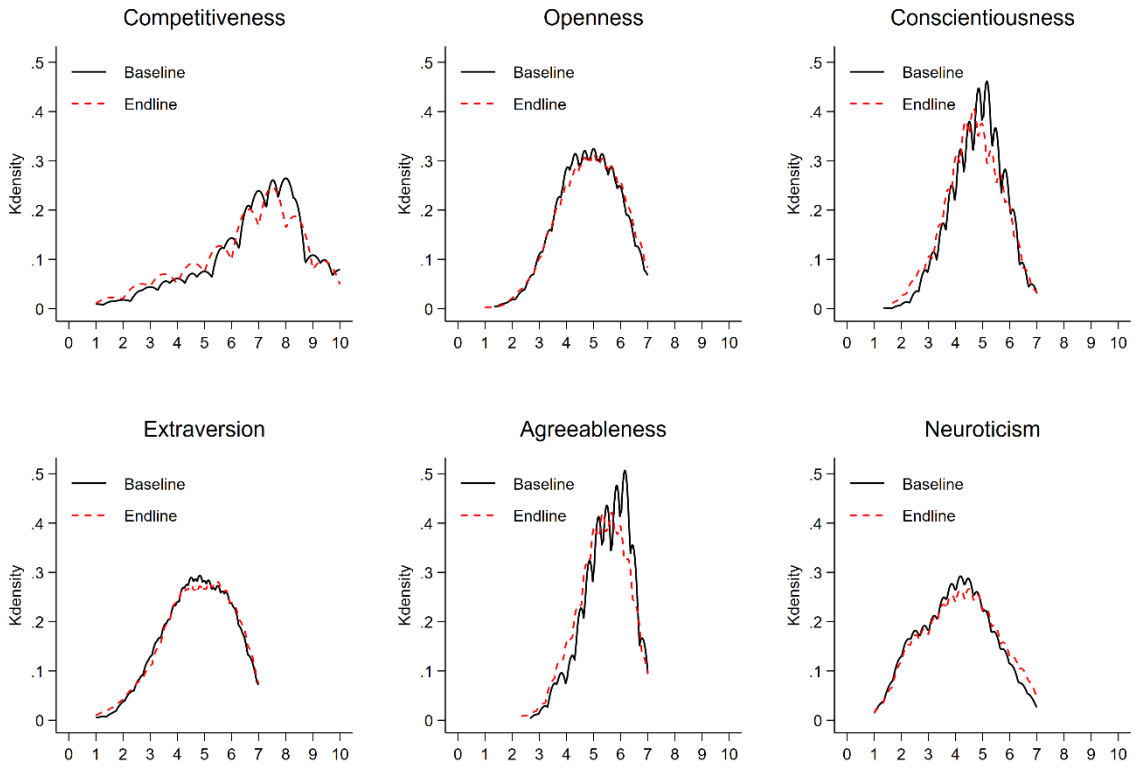


Notes: The figure shows the means and 95% confidence intervals for the Big Five personality traits separately for four samples: (1) the whole sample of respondents of the German Socio-Economic Panel (SOEP) in 2019 who were 18 years old or older (N = 25,979), (2) the young adult sample of the 2019 G-SOEP who were aged 18–22 (N = 1,920), (3) the whole sample of students in our setting who answered the baseline survey from 2018/2019 to 2023/2024 (N = 5,232); and (4) our final analysis sample of students (N = 1,229). Samples (2) and (4) are subsamples of (1) and (3), respectively. For our student sample, we plot summary statistics of their baseline personality. All respondents in the SOEP and our study reported the Big Five personality traits on the same scale (1–7 point Likert scale), so the raw scores are comparable across samples.

ONLINE APPENDIX

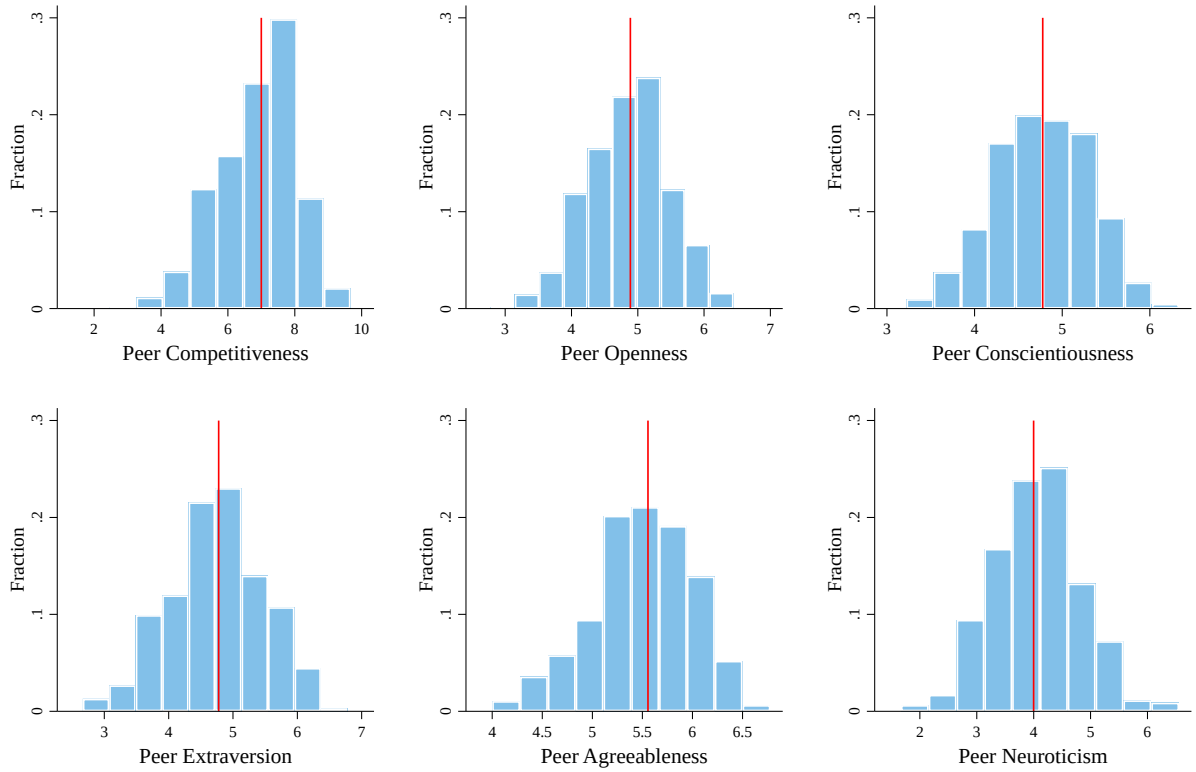
Appendix A: Additional Tables and Figures

Figure A1: The Distribution of Personality Traits at Baseline and Endline



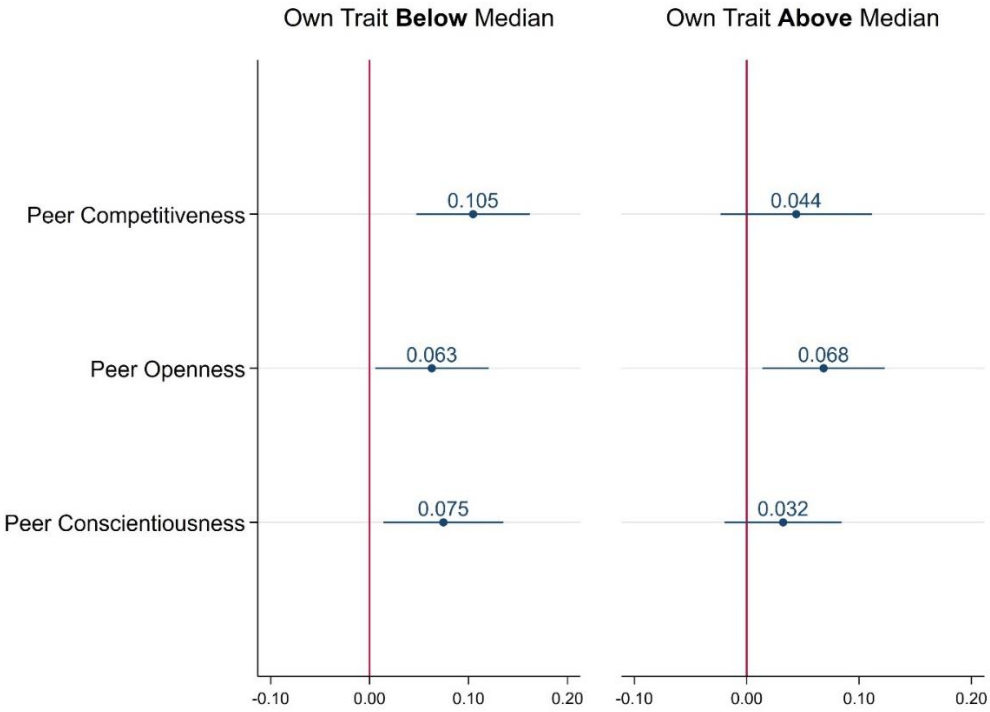
Notes: The figure shows the distribution of six personality traits at the baseline and endline (the raw scores) for the estimation sample (N = 1,229).

Figure A2: Histogram of Peer Personality Measures



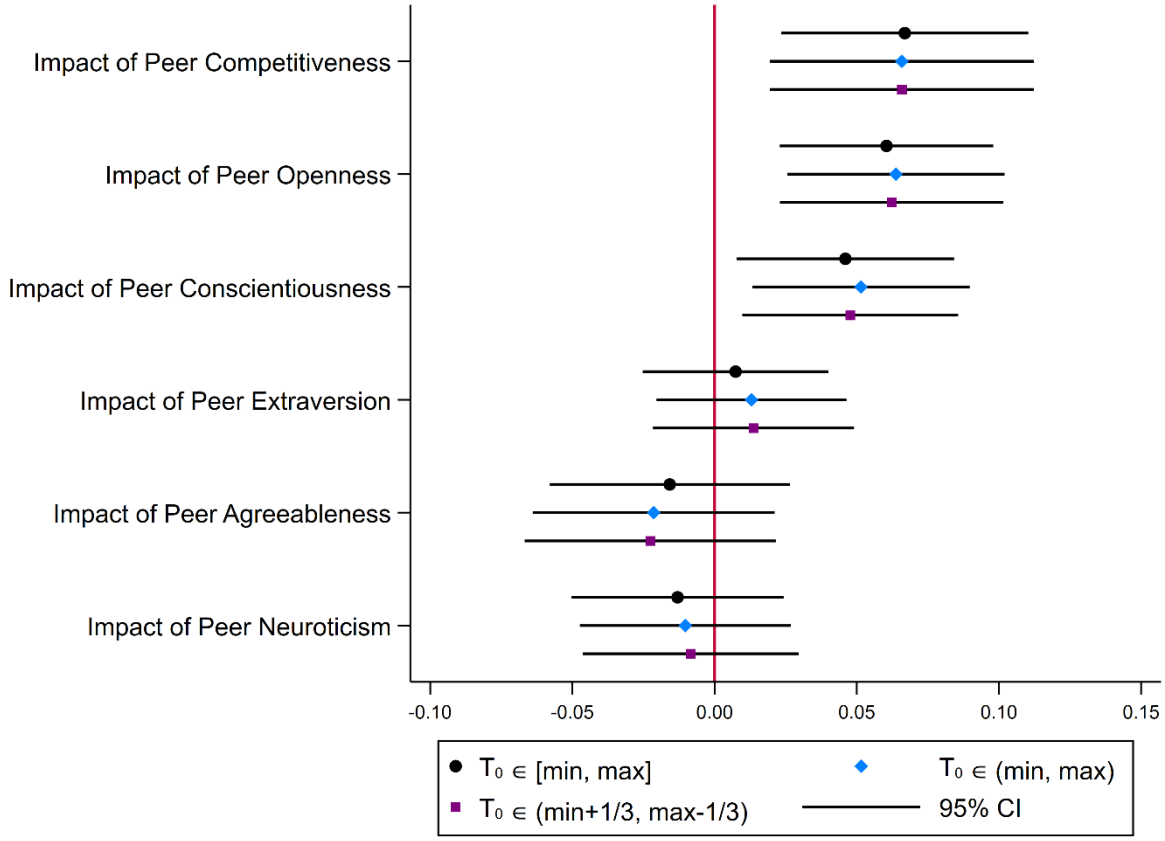
Notes: The figure shows the distribution of peer personality traits at baseline, using raw scores on a scale of 1 to 7 or 10. The vertical red line shows the median level of a peer trait.

Figure A3: Heterogeneous Effects: Own Trait Below Versus Above Median



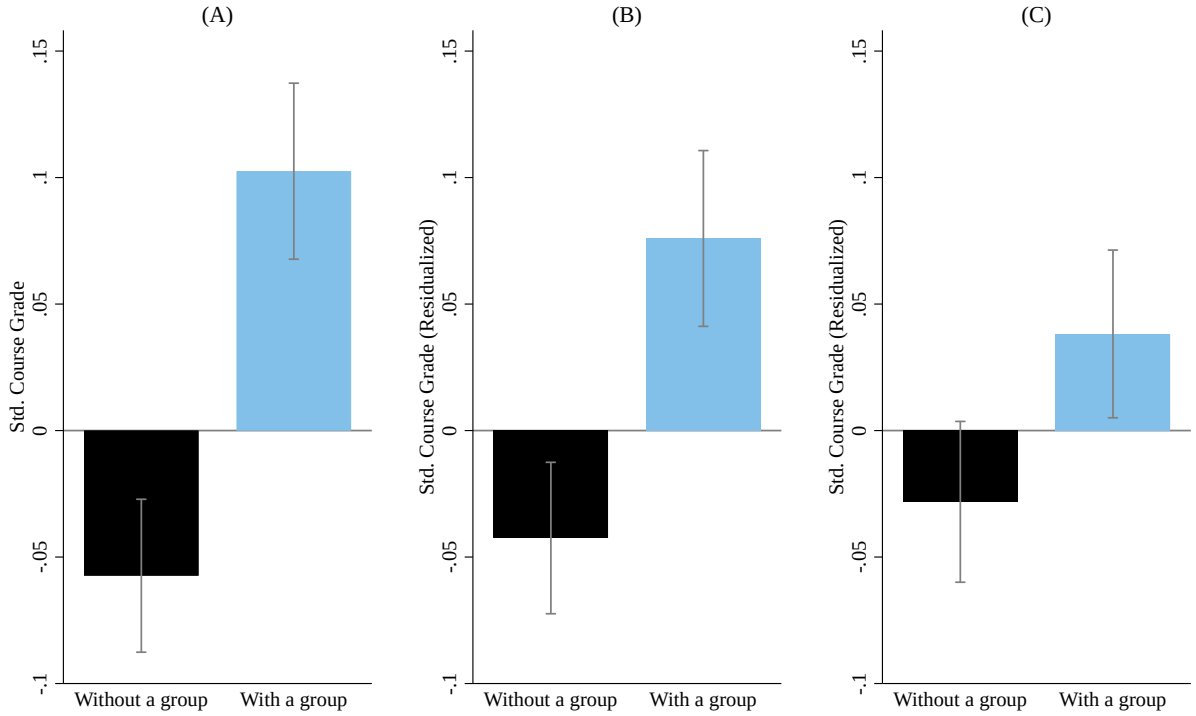
Notes: The figure shows heterogeneous effects by the level of own personality at baseline. For each trait, we divide students into two subgroups: those with below-median and above-median levels of the trait. Then we estimate the impact of the peer trait on own trait for the two subgroups, respectively. All regressions control for all individual characteristics at the baseline, as in column (3) of Table 4. None of the between-group differences in impacts reach significance at the 5% level. Error bars indicate 95% confidence intervals based on robust standard errors clustered at the study group level.

Figure A4: Robustness to Dropping Extreme Values of Personality



Notes: The figure shows the estimated spillover effects of six personality traits for the full sample, the sample of students whose baseline level of a trait is in the range of above the minimal value (1) and below the maximum value (7 or 10), and the sample of students whose baseline level of a trait is in the range of above the minimal value plus 1/3 and below the maximum value minus 1/3. Each point estimate is derived from one OLS regression using the same specification as in column (3) of Table 4. Error bars indicate 95% confidence intervals.

Figure A5: Course Grade for Students With or Without Study Groups



Notes: The figure shows the standardized final course grade by whether a student was in a study group. Panel (A) shows the raw pattern. Panel (B) shows the pattern of residualized grades after controlling for cohort and study program fixed effects. In Panel (C), we further control for all student characteristics listed in column (3) of Table 4, such as gender, age, high school grades, and personality traits, when residualizing the course grade. The error bars indicate 95% confidence intervals. The difference in (residualized) grades between students without study groups and students with groups is statistically significant at the 1% level in all panels.

Table A1: Baseline Characteristics by Group Registration

	Not Registered (N = 2,992)		Registered (N = 2,211)		Difference
	mean	*sd*	mean	*sd*	*p*-value
Competitiveness	6.779	1.987	6.879	1.999	0.074
Openness	4.781	1.105	4.867	1.084	0.005
Conscientiousness	4.837	0.952	4.790	0.964	0.078
Extraversion	4.676	1.263	4.758	1.241	0.019
Agreeableness	5.452	0.926	5.509	0.879	0.025
Neuroticism	4.074	1.297	4.043	1.309	0.397
High school math grade	4.537	0.810	4.568	0.793	0.171
High school language grade	4.763	0.553	4.755	0.595	0.593
Female	0.413	0.492	0.434	0.496	0.128

Notes: The table shows the summary statistics of the baseline characteristics by group registration and the *p*-values for tests of differences between students who are registered and students who are not.

Table A2: Test for Balancing and Selective Attrition:

Impact of Baseline Peer Personality on Own Personality and Observing Endline Personality

	(1) Competitiveness	(2) Openness	(3) Conscientiousness	(4) Extraversion	(5) Agreeableness	(6) Neuroticism
Panel A: Dependent Variable: Own Personality at Baseline						
Peer Competitiveness	0.005					
	(0.027)					
Peer Openness		-0.028				
		(0.026)				
Peer Conscientiousness			-0.053*			
			(0.029)			
Peer Extraversion				-0.032		
				(0.036)		
Peer Agreeableness					-0.050*	
					(0.030)	
Peer Neuroticism						0.039
						(0.028)
Observations	1,776	1,776	1,776	1,776	1,776	1,776
R-squared	0.343	0.299	0.312	0.277	0.190	0.300
Panel B: Dependent Variable: Endline Survey Response						
Peer Competitiveness	0.005					
	(0.012)					
Peer Openness		-0.002				
		(0.012)				
Peer Conscientiousness			-0.010			
			(0.011)			
Peer Extraversion				-0.006		
				(0.011)		
Peer Agreeableness					0.008	
					(0.012)	
Peer Neuroticism						-0.000
						(0.011)
Observations	1,776	1,776	1,776	1,776	1,776	1,776
R-squared	0.040	0.040	0.040	0.040	0.040	0.040

Notes: Panel A examines whether peer personality predicts own personality at the baseline; Panel B tests whether peer personality at the baseline affects observing own personality at the endline. Each column represents one OLS regression. The sample used for analysis is 1,776 students in 340 groups, where all students reported their baseline personality traits. Out of the 1,776 students, 1,229 students (69%) reported their endline personality. The independent variables of interest are standardized peer personality traits at the baseline. All models control for cohort-by-program fixed effects. Results in Panel B are very similar even if we control for all individual characteristics at the baseline. Panel A uses robust standard errors, and Panel B uses clustered standard errors (shown in parentheses).

Table A3: Pairwise Correlations Between Personality Traits

	(1) Competitiveness	(2) Openness	(3) Conscientiousness	(4) Extraversion	(5) Agreeableness	(6) Neuroticism
Competitiveness	1					
Openness	0.059 [0.125]	1				
Conscientiousness	0.280 [0.000]	0.066 [0.005]	1			
Extraversion	0.134 [0.000]	0.257 [0.000]	0.126 [0.000]	1		
Agreeableness	-0.069 [0.004]	0.067 [0.005]	0.180 [0.000]	-0.027 [0.254]	1	
Neuroticism	-0.176 [0.000]	-0.028 [0.243]	-0.088 [0.000]	-0.230 [0.000]	-0.004 [0.875]	1

Notes: Pairwise correlations at the baseline (n = 1,776). *p*-values in brackets.

Table A4: Cronbach's Alpha Coefficient of Reliability

	(1) Baseline	(2) Endline	(3) Follow-Up
Openness	0.615	0.623	0.742
Conscientiousness	0.583	0.611	0.761
Extraversion	0.809	0.812	0.827
Agreeableness	0.461	0.455	0.522
Neuroticism	0.780	0.806	0.796

Notes: Each cell reports the Cronbach alpha for each trait (measured with three items) at the baseline, endline, and follow-up survey. In the follow-up survey, we have six items measuring openness and conscientiousness; thus, the alpha coefficients improved substantially. When we use the same three items as in the baseline and endline, the alpha coefficients are 0.599 and 0.564 for openness and conscientiousness.

Table A5: Stability of Personality – Within-Trait Correlation Across Time

	(1) Competitiveness	(2) Openness	(3) Conscientiousness	(4) Extraversion	(5) Agreeableness	(6) Neuroticism
Baseline & Endline	0.640 [0.000]	0.770 [0.000]	0.690 [0.000]	0.804 [0.000]	0.687 [0.000]	0.758 [0.000]
Baseline & Follow-Up	0.620 [0.000]	0.705 [0.000]	0.634 [0.000]	0.793 [0.000]	0.651 [0.000]	0.731 [0.000]

Notes: The table shows the correlation between a trait's level at the baseline and its level at the endline or follow-up. *p*-values are in brackets. We use the sample from five cohorts to calculate the correlations between baseline and endline levels and the sample from the first three cohorts to calculate the correlations between baseline and follow-up levels.

Table A6: Main Results Excluding the COVID-19 Cohort (2020/21)

	(1)	(2)	(3)	(4)	(5)
Panel A: DV = Std. Own Competitiveness at Endline					
Std. Peer Competitiveness	0.064***	0.064***	0.066***	0.085***	0.083***
	(0.024)	(0.024)	(0.024)	(0.026)	(0.027)
	[0.009]	[0.008]	[0.007]	[0.001]	[0.002]
Panel B: DV = Std. Own Openness at Endline					
Std. Peer Openness	0.060***	0.061***	0.060***	0.064***	0.058***
	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)
	[0.004]	[0.004]	[0.005]	[0.003]	[0.009]
Panel C: DV = Std. Own Conscientiousness at Endline					
Std. Peer Conscientiousness	0.051**	0.051**	0.050**	0.053**	0.056**
	(0.021)	(0.021)	(0.021)	(0.024)	(0.025)
	[0.017]	[0.018]	[0.019]	[0.025]	[0.027]
Panel D: DV = Std. Own Extraversion at Endline					
Std. Peer Extraversion	0.017	0.015	0.013	0.018	0.018
	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)
	[0.336]	[0.415]	[0.477]	[0.355]	[0.359]
Panel E: DV = Std. Own Agreeableness at Endline					
Std. Peer Agreeableness	-0.001	-0.001	0.002	0.005	0.003
	(0.025)	(0.025)	(0.025)	(0.025)	(0.024)
	[0.965]	[0.959]	[0.927]	[0.843]	[0.891]
Panel F: DV = Std. Own Neuroticism at Endline					
Std. Peer Neuroticism	-0.019	-0.019	-0.019	-0.023	-0.031
	(0.021)	(0.021)	(0.021)	(0.022)	(0.023)
	[0.366]	[0.364]	[0.368]	[0.297]	[0.170]
Observations	957	957	957	957	957
Control Variables:					
Other own personality traits	N	Y	Y	Y	Y
Other own characteristics	N	N	Y	Y	Y
Other peer personality traits	N	N	N	Y	Y
Other peer characteristics	N	N	N	N	Y

Notes: Each column in each panel shows one coefficient from a separate OLS regression that uses the estimation sample excluding the cohort of 2020/21. All regressions include the baseline level of the dependent variable as well as randomization controls as defined in the empirical strategy. *Other own personality traits* include five other personality traits at the baseline, except for the one trait being examined in each panel. *Other own characteristics* include gender, business-economics major fixed effects, high school math and language grades, high school study hours, an indicator for German as the high school instruction language, course-retaking status, and age fixed effects. *Other peer personality traits* are five other peer personality traits at the baseline, except for the one being analyzed. *Other peer characteristics* include peers' gender, high school math and language grades, and whether German was the instruction language in high school. All dependent and independent variables are standardized. Robust standard errors clustered at the study group level are in parentheses. *p*-values are in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Leuven Test of Randomization

	F-statistic	*p*-value
Competitiveness	0.930	0.820
Openness	0.845	0.983
Conscientiousness	0.764	1.000
Extraversion	0.870	0.961
Agreeableness	0.867	0.964
Neuroticism	0.979	0.602
Female	0.753	1.000
High School Math Grade	0.996	0.515
High School Language Grade	1.097	0.113
High School Study Hours per Week	0.971	0.641

Notes: We regress each baseline characteristic on study program and cohort dummies and derive the residuals. Then, we regress the residuals on study group dummies and test the joint significance of group dummies. The table reports the *F*-statistic and *p*-value separately for each baseline characteristic.

Table A8: Balance Test: Joint Effects of Peer Personality on Own Personality at Baseline

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Own Personality Traits at Baseline					
	Competitive-ness	Open-ness	Conscientious-ness	Extraversion	Agreeable-ness	Neuroticism
Peer Competitiveness	-0.008	0.016	0.034	-0.007	0.039	0.018
	(0.029)	(0.026)	(0.026)	(0.025)	(0.025)	(0.027)
Peer Openness	0.045*	-0.038	0.004	0.030	0.012	-0.061**
	(0.026)	(0.031)	(0.028)	(0.027)	(0.028)	(0.026)
Peer Conscientiousness	0.005	-0.012	-0.072**	0.013	0.030	0.002
	(0.028)	(0.027)	(0.032)	(0.028)	(0.028)	(0.027)
Peer Extraversion	0.000	0.008	0.049*	-0.049	-0.006	0.012
	(0.028)	(0.029)	(0.028)	(0.037)	(0.028)	(0.026)
Peer Agreeableness	0.012	0.024	0.026	-0.017	-0.040	0.028
	(0.026)	(0.028)	(0.027)	(0.026)	(0.032)	(0.025)
Peer Neuroticism	0.003	-0.036	0.012	0.006	0.024	0.050
	(0.025)	(0.026)	(0.025)	(0.024)	(0.026)	(0.032)
Observations	1,229	1,229	1,229	1,229	1,229	1,229
R-squared	0.332	0.317	0.320	0.307	0.315	0.293
Test of joint significance of six peer personality traits						
F-statistic	0.608	0.867	1.537	0.494	1.245	1.466
p-value	0.724	0.519	0.164	0.813	0.282	0.188

Notes: The table estimates whether peer personality jointly predict own personality at baseline. Each column is derived from one regression that controls for program-by-cohort fixed effects and the leave-out-mean of each personality trait at the program-by-cohort level. Robust standard errors clustered at the study group level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Full Spillover Matrix with Multiple Testing Correction

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Own Personality Traits at Endline					
	Competitive-ness	Openness	Conscientious-ness	Extraversion	Agreeable-ness	Neuroticism
Peer Competitiveness	0.078***	0.004	-0.005	-0.048**	-0.017	-0.024
	(0.024)	(0.021)	(0.021)	(0.020)	(0.023)	(0.020)
	[0.001]	[0.865]	[0.800]	[0.016]	[0.464]	[0.233]
	{0.001}	{0.999}	{0.999}	{0.046}	{0.996}	{0.915}
Peer Openness	-0.054**	0.063***	-0.007	0.006	-0.021	0.019
	(0.023)	(0.020)	(0.022)	(0.018)	(0.022)	(0.020)
	[0.019]	[0.002]	[0.759]	[0.715]	[0.338]	[0.346]
	{0.061}	{0.002}	{0.999}	{0.999}	{0.984}	{0.984}
Peer Conscientiousness	-0.029	-0.005	0.056**	0.032	-0.025	0.013
	(0.024)	(0.020)	(0.023)	(0.020)	(0.025)	(0.021)
	[0.220]	[0.816]	[0.015]	[0.117]	[0.326]	[0.522]
	{0.902}	{0.999}	{0.044}	{0.634}	{0.983}	{0.998}
Peer Extraversion	0.025	-0.043**	-0.028	0.010	-0.004	-0.010
	(0.025)	(0.019)	(0.023)	(0.018)	(0.024)	(0.022)
	[0.324]	[0.029]	[0.212]	[0.579]	[0.857]	[0.627]
	{0.983}	{0.116}	{0.901}	{0.999}	{0.999}	{0.999}
Peer Agreeableness	0.011	0.027	-0.015	0.035**	-0.010	-0.021
	(0.021)	(0.021)	(0.020)	(0.017)	(0.021)	(0.022)
	[0.581]	[0.196]	[0.446]	[0.036]	[0.645]	[0.338]
	{0.999}	{0.879}	{0.996}	{0.157}	{0.999}	{0.984}
Peer Neuroticism	0.010	0.027	0.041*	0.031*	0.012	-0.019
	(0.025)	(0.020)	(0.023)	(0.018)	(0.023)	(0.021)
	[0.693]	[0.183]	[0.068]	[0.092]	[0.596]	[0.350]
	{0.999}	{0.860}	{0.360}	{0.508}	{0.999}	{0.984}
Test of the joint significance of six peer personality traits						
F-statistic	3.08	2.89	1.80	3.10	0.86	0.52
p-value	0.006	0.009	0.098	0.006	0.525	0.792

Notes: Each column represents one OLS regression in which the dependent variable is one of the six own personality traits at the endline (standardized). The independent variables of interest are six peer personality traits at the baseline (standardized). The specification of each regression is the same as in column (5) of Table 4. Robust standard errors clustered at the study group level are in parentheses, and original *p*-values are in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. To account for multiple hypothesis testing, we apply the Romano–Wolf correction with resampling procedures to all 36 estimates and show the corrected *p*-values in curly brackets. Specifically we resample group clusters within cohort and study program strata 10,000 times. The last two rows of the table show the *F*-statistics and *p*-values derived from the tests of joint significance of six peer personality traits.

Table A10: Impact of Peer Personality on Follow-Up Survey Response

Dependent Variable: Follow-Up Survey Response	(1)	(2)	(3)	(4)	(5)	(6)
Peer Competitiveness	0.004					
	(0.014)					
Peer Openness		0.005				
		(0.015)				
Peer Conscientiousness			-0.001			
			(0.015)			
Peer Extraversion				-0.022		
				(0.013)		
Peer Agreeableness					0.013	
					(0.014)	
Peer Neuroticism						-0.010
						(0.014)
Observations	1,856	1,856	1,856	1,856	1,856	1,856
R-squared	0.063	0.063	0.063	0.065	0.063	0.063

Notes: Each column represents one OLS regression in which the dependent variable is an indicator for whether a student reports personality in the follow-up survey in the summer of 2021 or 2022. For 2018/19 to 2020/21 cohorts, we create a two-wave panel dataset for responses in both follow-up surveys; for the 2021/22 cohort, we only examine their responses to the 2022 survey. All specifications control for stratum-by-cohort fixed effects and fixed effects for the follow-up year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: A Simple Model on Personality Change Under the Influence of Peers

Consider a student who starts university education with a vector of K personality traits $\mathbb{T}_0 = [T_0^1, T_0^2, T_0^3, \dots, T_0^K]$. Personality is malleable and the student can change their traits from the baseline level to a new level $\mathbb{T} = [T^1, T^2, T^3, \dots, T^K]$. For simplicity we assume all trait levels to be strictly positive: $T_0^k, T^k > 0, \forall k \in 1,2, \dots, K$. Changes in traits are costly and the student obtains utility from academic achievement. Personality traits may affect achievement.

Our key modeling assumption is that peers affect the *costs* of personality change, for example, through social learning or social pressure. In the following, we will derive students' optimal personality vector ($\mathbb{T}^*$) in the absence of peers. After that, we will derive students' optional personality when exposed to peers who have higher, lower, or the same trait level.

No peers: As a benchmark, we first examine optimal personality development without peer influences. The student chooses their optimal personality vector ($\mathbb{T}^*$) by solving the following maximization problem:

$$\max_{\mathbb{T}=[T^1,T^2,\dots,T^K]} F(\mathbb{T}) - C(\mathbb{T};\mathbb{T}_0) = \sum_{k=1}^{K} f(T^k) - c\left(T^k; T_0^k\right), \tag{2}$$

where $f(T^k)$ denotes the production function of academic achievement that depends on trait T^k. $c\left(T^k; T_0^k\right)$ denotes the costs of adjusting the trait from the baseline level T_0^k to T^k. And $f(T^k)$ allows for personality traits to have different effects on achievement.[21]

For simplicity, we assume that traits (T^k) are substitutes—academic achievement is a linear function of each trait—$F(\mathbb{T}) = \sum_{k=1}^{K} f(T^k) = \sum_{k=1}^{K} \alpha^k T^k$, where α^k denotes the marginal benefit of raising T^k. We label traits as *productive* if higher levels of these traits lead to higher achievement: $\alpha^k > 0$. Other traits may have a negative or no impact on academic achievement: $\alpha^k \leq 0$.

We assume that students find changing their personality costly (McCrae and Costa 1994). The costs of changing personality increase with the distance from the baseline level of the given trait:

$$\mathrm{c}\left(T^k; T_0^k\right) = \begin{cases} \left(T^k - T_0^k\right)^\gamma, & \text{if } T^k \geq T_0^k \\ \left(T_0^k - T^k\right)^\gamma, & \text{if } T^k < T_0^k, \end{cases} \tag{3}$$

[21] We can also generalize the utility function by including other outcomes that may depend on personality, such as mental health and social integration. The intuitions of our framework remain the same: some personality traits are overall helpful and generate positive returns, while other traits may generate negative net returns.

with $\gamma > 1$ capturing that it is increasingly costly to move further away from the initial level. Without loss of generality, we assume $\gamma = 2$ and that costs are identical for all traits. These simplifying assumptions about the cost structure and γ are not necessary to arrive at the model's qualitative predictions. In equilibrium, the optimal level of a personality trait is determined by marginal benefit and marginal cost:

$$\alpha^k = 2(T^{k*} - T_0^k),\ k \in \{1,2,\dots,K\}. \tag{4}$$

The left-hand side of the equation is the marginal benefit, and the right-hand side is the marginal cost. Equation (4) shows that the optimal level of a productive trait is always above the baseline level: $T^{k*} > T_0^k$. For traits that are not relevant for academic achievement, students' optimal strategy is to avoid any costly changes: $T^{k*} = T_0^k$.

Peer Influences: We next consider how the presence of peers affects personality development. Suppose that the student is exposed to one peer or a group of peers whose baseline personality also consists of a vector of traits $\mathbb{T}_p = \left[T_p^1, T_p^2, T_p^3, \dots, T_p^K\right]$. For each trait, peer levels can be lower, higher, or equal to the student's initial level: $T_p^k \gtreqless T_0^k$. Peers may affect the costs of molding personality by acting as an example, providing a reference point, or creating peer pressure. Through these mechanisms, it becomes less costly to converge toward, and more costly to deviate from, peer levels.

We assume that the cost function remains centered at T_0^k under the influence of peers—the costs are always the lowest and equal to zero when personality change is not initiated. This assumption highlights that changing personality is difficult and requires effort, and this aversion to change is stronger than the conforming effect of peers. Without loss of generality, we can formalize the cost function in the presence of peers as:

$$c\left(T^k; T_0^k, T_p^k\right) = \begin{cases} \frac{T_0^k}{T_p^k}\left(T^k - T_0^k\right)^2, & \text{if } T^k \geq T_0^k \\ \frac{T_p^k}{T_0^k}\left(T_0^k - T^k\right)^2, & \text{if } T^k < T_0^k. \end{cases} \tag{5}$$

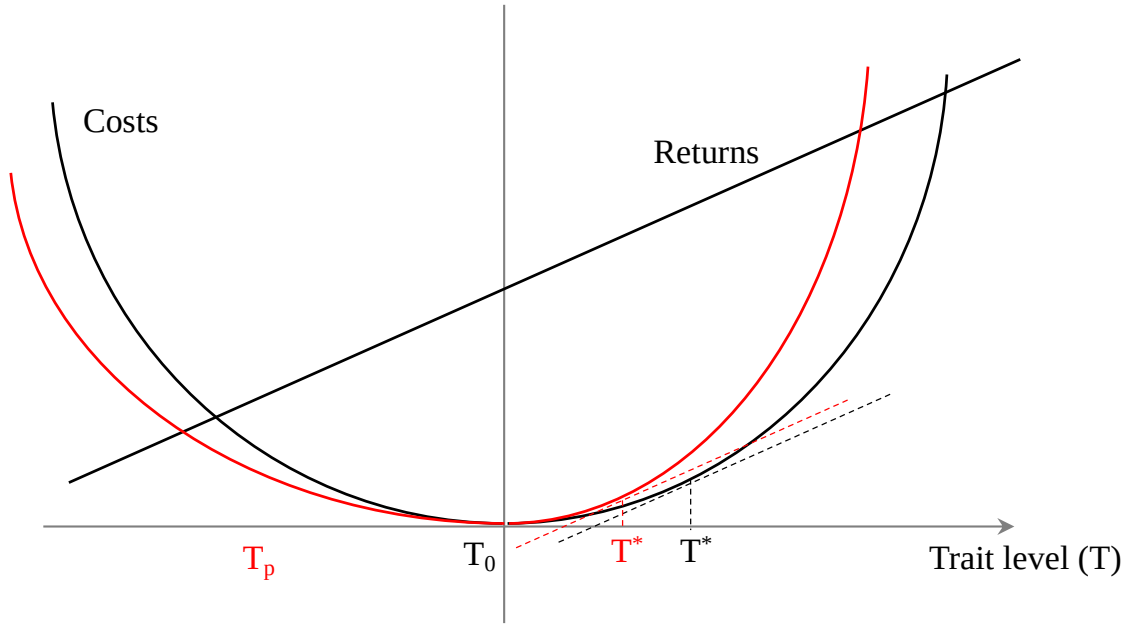
Equation (5) shows that when exposed to peers with the same trait level, $T_p^k = T_0^k$, costs are identical to the scenario without peers. When exposed to peers with higher levels of a trait, $T_p^k > T_0^k$, the costs of increasing (decreasing) own trait level are lower (higher) compared to the case without peers. Finally, we can derive the optimal trait levels as follows:

$$T^{k*}\left(T_0^k, T_p^k\right) = \begin{cases} \frac{T_p^k}{2T_0^k}\alpha^k + T_0^k, & \text{if } \alpha^k \geq 0 \\ \frac{T_0^k}{2T_p^k}\alpha^k + T_0^k, & \text{if } \alpha^k < 0. \end{cases} \tag{6}$$

Prediction: The key prediction of this framework is that for personality traits that do not affect achievement, where $\alpha^k = 0$, trait levels at the baseline are optimal (i.e., the student has no incentive to change personality) and there will be no personality spillovers. Peer personality only creates spillover effects for traits that affect academic achievement, that is, when α^k is unequal to zero. Put differently, optimal trait levels increase in peer levels if a trait is relevant for academic performance.

Figure A6 shows a stylized example of how the presence of peers affects the optimal trait level. In this example, the black curves represent the scenario without peers or when peers' level of a trait is equal to own trait level ($T_p^k = T_0^k$). The upward-sloping return curve means that the trait is productive for academic achievement. Thus, the student is incentivized to raise the trait level ($T^{k*} > T_0^k$). This academic motivation is somewhat offset by the presence of peers who have lower levels of this trait ($T_p^k < T_0^k$). Therefore, the optimal trait level decreases but remains above the baseline level.

Figure A6: Peers' Influence on Personality Development—An Example



Notes: T_0 denotes the baseline level of a personality trait, T_p denotes peers' baseline level of this trait, and T^* represents the optimal trait level. The *returns* curve is the academic output function, which increases linearly with the trait level. The *costs* curve is the cost function of personality change. The optimal trait level is reached when the marginal cost equals the marginal return. The example shows that when exposed to peers with a lower trait level, the optimal own trait level decreases. Similarly, when exposed to peers with a higher trait level, the optimal own trait level increases.

Appendix C: Selected Questions from the Baseline/Endline Survey

Table A11: Selected Questions from the Baseline Survey

Question	Answering options
What is your gender?	1) Male; 2) Female
What is your age?	1) 18 and below; 2) 19; 3) 20; 4) 21; 5) 22 and above
What was the main language of instruction in your high school?	1) German; 2) Italian; 3) French; 4) English; 5) Spanish; 6) Other
In high school, how many hours per week did you typically study for school on your own?	Please fill in a number
In high school, how many hours per week did you typically study for school interacting with others?	Please fill in a number
Below are different qualities a person can have. To what degree do they apply to you?	
I see myself as someone who does a thorough job	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is communicative, talkative	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is sometimes somewhat rude to others	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is original, comes up with new ideas	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who worries a lot	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who has a forgiving nature	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who tends to be lazy	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is outgoing, sociable	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who values artistic, aesthetic experiences	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who gets nervous easily	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who does things effectively and efficiently	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is reserved	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is considerate and kind to others	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who has an active imagination	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is relaxed, handles stress well	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
In general, how competitive do you consider yourself?	Scale: 1 = "not competitive at all" to 10 = "very competitive"

Table A12: Selected Questions from the Endline Survey

Question	Answering options
How often have you studied for "Mikroökonomik 1" together with each of your study mates? For example, you can attend tutorials together, work on problem sets and old exams, go through lecture slides together, etc.	
[Name peer 1] ; [Name peer 2] ; [Name peer 3]	0) Never; 1) less than once per month; 2) once per month; 3) 2 to 3 times per month; 4) once a week; 5) multiple times per week
How often have you joined social events together with each of your study mates? For example, you can do sports, watch movies, have drinks and dinner together, etc.	
[Name peer 1] ; [Name peer 2] ; [Name peer 3]	0) Never; 1) less than once per month; 2) once per month; 3) 2 to 3 times per month; 4) once a week; 5) multiple times per week
Except for the students of the assigned study group, have you also studied with other students for "Mikroökonomik 1"?	1) Yes; 2) No, I have not studied with other students
Below are different qualities a person can have. To what degree do they apply to you?	
I see myself as someone who does a thorough job	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is communicative, talkative	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is sometimes somewhat rude to others	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is original, comes up with new ideas	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who worries a lot	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who has a forgiving nature	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who tends to be lazy	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is outgoing, sociable	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who values artistic, aesthetic experiences	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who gets nervous easily	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who does things effectively and efficiently	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is reserved	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is considerate and kind to others	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who has an active imagination	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
I see myself as someone who is relaxed, handles stress well	Scale: 1 = "does not apply to me at all" to 7 = "applies to me perfectly"
In general, how competitive do you consider yourself?	Scale: 1 = "not competitive at all" to 10 = "very competitive"