Peer creativity and academic achievement

Lent, M. van

Economics of Education Review 106 (2025) 102649

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Peer creativity and academic achievement☆

Max van Lent

Department of Economics at Leiden University, The Netherlands
IZA, Germany

ARTICLE INFO

JEL classification:
I21
I24
J24

Keywords:
Peer effects
Academic achievement
Creativity
Field experiment

ABSTRACT

This paper studies the relationship between the creative abilities of study peers and academic achievement. We conduct a novel large scale field experiment at university, where students are randomized into work groups based on their score on a creativity test prior to university entry. We show that the creative abilities of peers matter for a student's academic achievement. A one standard deviation higher creativity peer group improves study performance by 6.2 to 7.6 percentage points. Further analysis suggests that students exposed to creative peers become more creative, but do not adjust their overall study effort. This is in line with the idea that creative approaches and questions from peers help students master the study material better. Overall, our study highlights the importance of peer effects of creative students in shaping academic outcomes.

1. Introduction

Creativity – defined as the ability to produce novel ideas or solutions that are useful and appropriate in a given situation (Amabile, Conti, Coon, Lazenby, & Herron, 1996; Bradler, Neckermann, & Warnke, 2019) – is seen as an important skill that is essential to problem solving and (for now) mostly resistant to automation. It is important for entrepreneurship and innovation (Erat & Gneezy, 2016; Gross, 2020) and a driver of the economy (Charness & Grieco, 2019). On an individual level, creativity has significant payoffs in terms of educational attainment and labor market outcomes (Gill & Prowse, 2024).

There is a by now large literature showing that peer effects are an important driver of educational performance, see e.g. Sacerdote (2011) for a literature review as well as a discussion at the end of this introduction. While the importance of creativity is by now well studied, the impact of peer creativity in education (and more general) has received surprisingly little attention.

In this paper we study the impact of the creative abilities of randomly assigned study peers on students' academic achievement. In this context, before entering university, students participate in a survey. As part of this research, we include two creativity tests to this survey: the Remote Associates Test (RAT) and a domain specific version of the Kaufman Domains of Creativity Scale (K-DOCS). Based on students' score on the RAT we assign them randomly to their own creativity type group (high or low) or to a mixed creativity group. We choose the RAT as our main measure of creativity for several reasons, which are carefully explained in Section 3. The most important reasons are that the RAT is specifically developed to measure creativity (see Mednick (1963)) and the RAT is shown to strongly correlate with other creativity measures (see e.g. Pesout and Nietfeld (2021)). We subsequently estimate how peer creativity impacts academic achievement. The creative abilities of study peers can positively influence students in at least three ways. First, students who work together with creative peers can learn from the questions that creative peers ask or from discussions with creative peers, and learn strategies and techniques that help them increase their study performance. Second, students may become more motivated and engaged from experiencing creativity from their peers, for instance when creative peers make learning more fun and challenging. This leads students to increase their study effort and as a consequence, their performance increases. Finally, there may be a direct effect from creative students who stimulate their peers to think outside the box, resulting in improved creative ability and as a consequence, improved study performance. On the contrary, the creative abilities of study peers may also negatively impact students. The novel ideas and questions of creative students may lead other students (especially the less creative ones) to be challenged too much, leading to confusion and discouragement. This may lead students to reduce effort. In addition, teachers may adjust their teaching methods as a consequence of the type of questions and discussions that emerge in class. This may also impact students' study performance.

Our results show that students' creativity is positively associated with their own study performance (GPA). A one standard deviation increase in a student's own creativity is associated with an increase of 8.3

☆ I thank participants at MBEES 2022, NPS 2022, COPE 2023, the WiP seminar Economics in Leiden, and the Field Days 2024 for their valuable comments and suggestions. This RCT is registered as AEARCTR-0012099 (van Lent, 2023).

E-mail address: m.van.lent@law.leidenuniv.nl.

https://doi.org/10.1016/j.econedurev.2025.102649
Received 25 October 2024; Received in revised form 13 March 2025; Accepted 13 March 2025
Available online 5 April 2025

percentage points in GPA. This finding is in line with the literature, see e.g. Naderi, Abdullah, Aizan, Sharir, and Kumar (2009, 2010), Kaufman (2010), Mourgues, Tan, Hein, Elliott, and Grigorenko (2016), Atwood and Pretz (2016), and Zhang, Ren, and Deng (2020) who report either insignificant or small and positive associations between creativity and academic achievement. The key contribution of this paper lies in the ability to study the effect of peer creativity on study achievement. We find that students benefit from being assigned to higher creativity peer groups. Specifically, high creativity students perform better in high creativity groups (as compared to mixed creativity groups) and low creativity students seem to benefit from mixed groups (as compared to low creativity groups), though this latter finding is insignificant. When we study the actual mean of peer creativity we find that a one standard deviation increase in the mean of peer creativity increases study performance by 11.4 percentage points of a standard deviation (without control variables) to 6.2 percentage points of a standard deviation when extensively controlling for individual as well as peer characteristics.

As a robustness test, we replicate our main findings using an alternative measure of creativity, a stated-creativity test: the K-DOCS. There is a small positive (but insignificant) effect of one's own creativity on study performance, and a one standard deviation increase in the score on the K-DOCS leads to a mostly insignificant increase in GPA of 1.1 to 8.7 percentage points. The fact that a revealed creativity test (such as the RAT) contains a much stronger effect than a stated creativity test (such as the K-DOCS) is in line with earlier findings from the literature.

With the aim of shedding light on what mechanisms drive the positive impact of creative peers on academic achievement, we use a follow-up survey six months after the randomization. These data show that students with creative peers become more creative, but they do not increase the amount of study hours. This implies that students become more productive per hour. Perhaps surprisingly, students are naive about the positive impact of creative peers. Students in higher creativity peer groups do not believe they learn more from their work group peers, as compared to students from lower creativity groups, nor do they allocate more hours towards studying with their peers.

This paper relates to two main strands of literature. The literature that studies the relationship between creativity and academic achievement and the literature on peer effects in education. Regarding the first strand of literature, there are many papers that study the association between creativity and academic achievement. See e.g. Gajda, Karwowski, and Beghetto (2017) for a recent literature overview of 120 papers from the 1960s onwards. On average, the correlation between creativity and academic achievement is 0.22, which is stable over time, and weaker for GPA measures than for measures using standardized tests. This is in the same direction, though somewhat larger than the correlation that we find.[1] Notably, the correlation is stronger for revealed creativity tests as compared to self-reported measures, which is also in line with our data.

There is a rich body of research studying peer effects in education, see Sacerdote (2011) for an extensive review of the literature. This literature is concerned with the estimation of causal effects of social spillovers within groups. It shows that peer test scores affect students.[2] Also gender and race composition of groups impact students.[3] So far there is surprisingly little research on the peer effects of non-cognitive skills and preferences. There are only a few recent exceptions, e.g. Golsteyn, Non, and Zölitz (2021) on the peer effects on persistence and risk preferences, Hancock and Hill (2022) and Shure (2021) on the peer effect of conscientious peers, Zárate (2023) who study the impact of peers' social skills on academic achievement, Shan and Zölitz (2022) on the impact of the big five of peers on academic achievement, and Zou (2024) on the impact of peers' persistence. The contribution of this paper to this literature is twofold. First, we are the first to study the effect of peer creativity on academic achievement. Second, by randomizing creativity block based into three groups – low creativity, high creativity, and mixed creativity – we increase the variation of peer creativity between groups. This approach is in line with Booij, Leuven, and Oosterbeek (2017) and Zárate (2023). As a consequence this paper avoids the issue that the exogeneous variation between groups is so limited that the estimates become imprecise and amplify bias, see also Angrist (2014) for an extensive discussion on this issue.

A challenge in estimating the impact of peer creativity – or any other personality trait or characteristic – on academic achievement is the fact that peer creativity can be related to other characteristics of peers that directly impact academic achievement. In order to mitigate this concern, we control extensively for peer characteristics that are known to affect study performance (and that may be correlated with creativity). We control for high school study grades, risk preferences, and persistence of peers in some of our regressions.

Our key finding – of positive creativity peer effects – has several implications. First, since creative students create positive spillovers, creativity can (and should) be used as a selection criteria into selective universities. Second, the positive peer effects imply that the value of creativity extends beyond the individual, this suggests that courses that stimulate creativity are more valuable than previously thought. Third, the positive spillovers that come from creative peers depend on the exposure of students to their peers. This implies that information provision to students about the benefits of creative peers may help them to reap the full benefits from peer creativity.

This paper proceeds as follows. In the next section we discuss the institutional setting. Section 3 describes the data, which includes both survey and administrative data and descriptive statistics. Section 4 explains the experimental design, Section 5 the empirical methodology, and Section 6 the results. The final section provides a discussion of the results and concludes.

2. Setting

This experiment took place in all full-time undergraduate study programs of a large Law School in The Netherlands. Law studies are consistently among the top 10 largest and most popular studies, and Law Schools in The Netherlands are not very selective. Students from all pre-university tracks can enter Law school. This particular Law School has an influx of around 1200 students per year who study Law or Criminology.

All first year students are assigned to small work groups of 25 to 30 students, before the start of each academic year.[4] Students spend their first days at university in these groups, learning about each other, the study program, and the information and communication systems they will be using during their studies. After these introduction days students participate in these same small groups for most of their study time for the rest of their first year of studies. The assignment of students to these work groups is binding. Students cannot switch to other groups.

Most courses consist of one or two main lecturers, who teach all students at once, and work groups where students learn together and work through cases, cooperate on assignments, and solve problem sets. The lectures are not designed to be very interactive, they are recorded for students to watch at their preferred time, and hence many students

[1] We find a correlation between creativity and a student's GPA of 0.134.

[2] See e.g. Carrell, Fullerton, and West (2009), Carrell, Sacerdote, and West (2013), De Giorgi and Pellizzari (2013), Duflo, Dupas, and Kremer (2011), Feld and Zölitz (2017), Lyle (2009), Whitmore (2005), Zimmerman David (2003).

[3] See e.g. Hoxby (2000), Angrist and Lang (2004), Hoxby and Weingarth (2005), Lavy and Schlosser (2011), Oosterbeek and van Ewijk (2014), and Gong, Lu, and Song (2021).

[4] The cohort 2020–2021 started during the COVID-19 pandemic. As a consequence, work group were made smaller and hosted most of the time online.

do not attend the lectures. The work groups on the other hand are interactive, and participation is mandatory. These work groups are designed with the purpose of studying together and learning from each other, as well as for the social aspect of studying at university. For instance, students occasionally have social activities with these groups. In other words, this work group system is designed with the purpose of helping students to be positively affected by each other. As a consequence of this institutional setting – with big lectures that are neither interactive nor compulsory and with small work groups that are interactive and mandatory – it is likely to have spillovers at the work group level, but not between different work groups.

In the first year students take 10 courses, depending on their study specialization.[5] Roughly 30% of their scheduled classes are lectures, the remaining 70% percent of their contact hours is in the smaller work groups.

In all study programs in our sample, students are graded on an absolute – instead of relative – grading scale. On top of that, work group teachers typically do not grade their own students. Instead, all teachers grade part of the exams of all students. These characteristics of the program avoid the potential issue that if all students in a work group perform better – for instance because of peer creativity – the grades are adjusted downwards because of grading on a curve. In addition, since exams are mostly graded by other teachers – and are mostly graded anonymously – there is hardly any room for favoritism or discrimination.

Students are allowed to initially fail a few courses in their first year of studies. They can progress to the second year when they complete at least 40 of the 60 course credits in the first year. In the second year students need to finish at least the remaining credits from the first year. Students who fail to obtain at least 40 course credits at the end of the first year have to drop out of the program.

3. Data

We have data on all full-time first year students of the cohorts 2019–2020, 2020–2021, and 2021–2022 of a Law School at a Dutch University from two different sources. The administrative data on all students that start in these undergraduate study programs include gender, high school grades, study track specialization, and study grades obtained at university. Second, we have data from two surveys. One that was administered before students entered university and were assigned to work groups, and a second survey that was administered during the second semester, i.e. six months later. These surveys were conducted online, sent from the office that also processes the student's paper work, and were part of the university's general intake procedure. Students fill in the first survey alone, independently, and before they meet other students. The surveys were matched with the administrative data. Both surveys contain questions on students' creativity. Specifically, we conduct two well-known creativity tests, the RAT and the K-DOCS. In addition, the second survey is supplemented with questions on study behaviors and work group dynamics. For more detail on the data collection and protocol, see Appendix C.

3.1. Remote associates test

There are many different tests used in the literature to measure creativity, without a consensus on what the best measure is, see e.g. Freund and Holling (2008) and Kaufman, Plucker, and Baer (2008) for such discussions. In this paper, the main measure of students' creative abilities is the Remote Associates Test (RAT).

The RAT (developed by Mednick (1962)) is a type of cognitive task that measures a person's ability to generate creative solutions by making connections between seemingly unrelated words. During the RAT, participants are presented three words, and are instructed to find a fourth word that is most strongly associated with all three words. The three words may seem unrelated at first, but there is a hidden link that can be discovered through creative thinking (Backman & Tuckman, 1972). The RAT is often used as a measure of creativity and problem-solving ability because it requires people to provide novel solutions, 'think outside the box', and make connections between concepts that seem unrelated. Several of these three word problems together form a measure of someone's creative abilities.[6]

We chose the Remote Associates Test (RAT) as our main creativity measure for several reasons. The test is developed with the intention to measure creativity specifically (Mednick, 1963) and validated by Chermahini, Hickendorff, and Hommel (2012) in the Dutch language. It is a language based creativity test, that can be easily incorporated in a survey. As Gajda et al. (2017) point out, linguistic creativity tests are more strongly associated with academic achievement as compared to figurative tests. Further, the test is a convergent thinking test, which correlates more strongly with academic performance (as compared to a divergent thinking test), see Yang and Zhao (2021). Finally, the RAT has a significant correlation with many other creativity measures. For instance, Pesout and Nietfeld (2021) show in a large sample of college students a significant correlation between the RAT and the product improvement test (PIT, as part of the Torrance Test of Creative Thinking, Torrance (1974)) and the Similarities Test (Walach and Kogan 1965). In Appendix D I provide more details about how the RAT is exactly conducted, and add a couple of example questions.

3.2. K-DOCS

Supplementary to the RAT, we conducted a stated creativity test. One benefit of using this additional test of creativity is that the stated test can be made domain specific. Secondly, for creativity spillovers someone's own assessment of their creativity may matter in particular, because it may affect the way in which students choose to interact with their peers more. For instance, students who believe they are creative may be more outspoken in group discussions. We use the Kaufman Domains of Creativity Scale (K-DOCS) as a starting point (see Kaufman (2012)) and adjust some of these questions to be specific of the academic study domain. In the end we developed and used the following statements: *Compared to others, how creative would you rate yourself for each of the following acts?* 1 = much less creative, 2 = less creative, 3 = neither more or less creative, 4 = more creative, 5 = much more creative.

1. Choosing the best solution to a study problem.
2. Helping other people cope with a difficult situation.
3. Thinking of many different solution to a study problem.
4. Helping other students during a difficult problem or assignment.
5. Thinking of new ways to help people with their studies.
6. Being able to offer constructive feedback based on my own reading of a paper.
7. Coming up with a new way to think about an old debate.
8. Debating a controversial topic from my own perspective.

In order to get an impression about the quality of the answers to these questions, we added a question in the end where the respondent is asked to rate how difficult they found it to answer the series of statements about their own creativity. A large majority (nearly 70%) of respondents state that they did not find it difficult (at all) to answer these questions.

[5] These seven specializations are: Dutch Law, Fiscal Law, Notarial Law, Law and Economics, Law and Entrepreneurship, International Business Law, and Criminology. In the first year all but one course are the same of the first six specializations. The criminology track is more different.

[6] A whole question bank can be found at: https://www.remote-associates-test.com/.

3.3. Descriptive statistics

This paper uses data from three cohorts of first-year students who have subscribed to a full-time undergraduate degree at a large Dutch Law School. The students are enrolled in one of the following seven programs: Dutch Law, Fiscal Law, Notarial Law, Law and Economics, Law and Entrepreneurship, International Business Law, or criminology. Dutch law is the largest with around 750 students a year. The other programs vary yearly between 50 and 125 students. The target group of students in this paper are 3682 students, 1090 in the 19–20 cohort, 1296 in the 20–21 cohort, and 1296 in the 21–22 cohort. These students are divided over 159 work groups in total. There are 1973 students who complete the survey. These 1973 are the 'full sample' throughout this paper, because having a creativity score (that students can only obtain by filling in the survey) is essential. The non-responding students are subsequently allocated randomly and equally to each group. More details about the assignment procedure will be provided in the next section.

In the Netherlands, Law schools typically have a majority of female students, in this setting 65% of the students enrolled is female. Students enroll into a study program after completing one of the pre-university high school tracks. In the cohorts starting in 2019, 2020, and 2021 at university the fraction of female pre-university high school graduates varies between 52% and 53.6%, see CBS (2023). The university keeps track of students' high school grades for Dutch language and Mathematics. Grades vary between 1 and 10, where 5.5 is the threshold of passing a course. Students in our sample score on average (for the 3 cohorts) 6.65 for Math and 6.72 for the Dutch language. The national averages for these same cohorts we see Math grades that vary between 6.70 and 6.83, and Dutch grades that vary between 6.47 and 6.56, see CBS (2023). This shows that students in our Law school score better in the Dutch language and worse in Mathematics. Since Law school is language intensive and requires relatively less math skills, we can interpret this finding as selection based on ability.

On the RAT students can score 0–12 points, 1 point per question. The average score in our sample is 5.82, which is similar to the score in Lee, Huggins, and Therriault (2014) who also administer the RAT among a sample of undergraduate students. Across the cohorts the difficulty of the questions was the same, and consequently their scores were also similar. Female students receive higher scores on the RAT which is in line with earlier studies.

The self-reported creativity test (K-DOCS) consisted of eight questions, where 1 indicated the lowest and 5 the highest score. Therefore the scores vary between 8 and 40. With an average of 29.58 students believe they are slightly more creative than their peers (a score of 24 would mean that students believe they are at the average of creativity in their groups). Males score higher on the K-DOCS.

We measure risk preferences using the response to the question: "In general, how willing are you to take risks?" For persistence we took six statements from the CP-SRLI that the respondent rates one a 5-point scale, ranging from never to always, see also Vandevelde, Van Keer, and Rosseel (2013) for a description.[7]

The last panel of Table 1 shows students' study performance at university. It is important to note that students are graded on an absolute scale, not relative to other students in their working group or cohort. GPA is the students' grade point average (between 1 and 10, where 5.5 is the threshold for passing a course). The next row gives the number of courses passed from the total of 10 courses. The GPA and number of courses completed is relatively low, because the group of students that drops out of the program is relatively large. In addition, students need to pass at least two thirds of their courses in order to proceed to the second year of their study program. We see that one out of five students drops out of their study program during the first year. All these study performance patterns are similar across cohorts. In Table A.1 we compare the full sample of our study with the students who do not participate in the first survey. These students have slightly lower high school grades, are more often male, and they perform worse at university. Recall however, that these students are equally distributed across all work groups, and therefore are not likely to impact the peer effects differently across groups.

Table 1
Descriptive statistics.

	Full sample	Males	Females
Male (%)	35.48	X	X
HS Grade Dutch language	6.72	6.58	6.79
	(0.68)	(0.73)	0(.64)
HS Grade Mathematics	6.65	6.63	6.65
	(1.02)	(0.94)	(1.06)
Creativity	5.82	5.61	5.94
	(2.12)	(2.07)	(2.14)
Stated Creativity	29.58	30.16	29.26
	(3.55)	(3.61)	(3.47)
Persistence	23.04	22.88	23.13
	(2.76)	(2.90)	(2.68)
Risk	6.96	7.30	6.77
	(3.55)	(1.46)	(1.35)
GPA	5.15	4.84	5.32
	(1.95)	(2.03)	(1.89)
Course credits	5.46	4.81	5.81
	(3.55)	(3.52)	(3.52)
Drop out (%)	20.1	23.7	18.1
Observations	1973	700	1273

The full sample refers to the sample of students how participated in the RAT (creativity test). this response is needed in order to estimate our individual measure of creativity. The standard deviations are in parentheses.

4. Experimental design

The aim of this study is to measure the impact of the creative ability of work group peers on students' academic achievement. We therefore elicit students' creative ability before entering university, and manipulate the assignment of students to peer groups. The timing for each of the three academic years is as follows:

In order to study how peer creativity affects students' performance, we can do better than only relying on natural variation stemming from a full randomization of students into work groups. This is the case because natural variation is limited and as a consequence the difference between high and low creativity groups stemming from only natural variation would be quite limited. This is especially the case if one wants to learn whether students should be assigned to similarly creative students or mixed across the entire creativity distribution. Such a recommendation is unlikely to follow from a fully randomized design, because groups of only high or only low creativity are expected to be underrepresented in such designs, see e.g. Angrist (2014) and Booij et al. (2017) for a thorough discussion. We therefore create additional variation using the experimental design. We do this in two steps. First, we assign students a type based on their creativity score. Second, based on this type we randomly allocate students to different group types with low, high, or mixed creativity peers. We next provide more detail about each of these steps.

[7] The statements are: (1) Even if I would rather do other things, I make myself start my schoolwork. (2) Even if my schoolwork is difficult or boring, I do my best. (3) Even if I would rather do other things, I finish my schoolwork. (4) I carry on until I finish my schoolwork. (5) During my schoolwork, I work attentively and do not take my mind off it. (6) If I am distracted while doing my schoolwork, I immediately try to continue working.

Month:	7	8	9	10	11	12	1	2	3	4	5	6	7
Randomization:		X											
Survey:	X	X X						X	X				
Period:	Holiday		Period 1		Period 2		Resits	Period 3		Period 4		Resits	

Fig. 1. Timeline of survey distribution, randomization, and course taking.
Note: Months refer to calendar months. For example, the first survey starts at the second half of July and Period 1 starts in the first week of September.

Table 2
Descriptive statistics: creativity types.

	High creativity	Low creativity	p-value: diff
Male (%)	35.10	38.75	0.010***
HS Grade Dutch language	6.77	6.64	0.000***
	(0.69)	(0.66)	
HS Grade Mathematics	6.66	6.63	0.612
	(1.05)	(0.97)	
Persistence	22.95	23.16	0.107
	(2.67)	(2.87)	
Risk preference	6.95	6.98	0.589
	(1.34)	(1.51)	
Study spec: Crim (%)	15.06	12.88	0.169
Fiscal (%)	3.59	8.06	0.000***
Notarial (%)	5.87	9.51	0.002***
Entrepreneurship (%)	10.95	10.59	0.801
International Bus. (%)	8.49	10.35	0.161
Economics (%)	5.08	3.61	0.119
Mono discipline (%)	50.96	45.01	0.009***
Observations	1142	831	

Note: The full sample refers to the sample of students who participated in the RAT (creativity test). High creativity refers to a score of 6 or higher (max score is 12) and low creativity refers to a score of at most 5. The standard deviations are in parenthesis.

4.1. Step 1: assign students a type

We assign students to one of three types based on their RAT score: high creativity, low creativity, or non-participant. The non-participant type consists of all students who had not participated in the survey at the moment of randomization. Some of those students still participate in the survey after the randomization but before the start of the academic year (when they first meet their work group students), these students still receive a creativity score. Those students who have completed the survey before randomization have a creativity score - a discrete score between 0 and 12 — and are allocated a type based on their score. The mean scores of participants were in each year between 5 and 6. Therefore, students with a creativity score of 5 or lower are assigned the low creativity type, and those with a score of 6 or higher are assigned the high creativity type.

Table 2 shows a comparison of the characteristics of low and high creativity students. We find minor differences in gender and high school grades in Dutch. High creativity students are more often female and score on average 0.13 better of their high school Dutch exams. In our regression models we control for these covariates.

4.2. Step 2: randomize students into work groups conditional on their type

Based on students' type (low or high creativity) we assign each student to work groups that are of three possible types: low creativity groups consisting of students who score low on creativity; high creativity consisting of students who score high on creativity, and mixed creativity groups; consisting of a mix between low and high creativity students. In mixed groups I created a balance between high and low creativity students under the constrained that students should be allocated to their own study specialization, and that there are low, high, and mixed groups (when there are at least three groups) per specialization. This let to mixed groups in some specializations and years to have at most 2 more high creativity types than in other mixed groups. All the students that did not complete the survey before the time of work group assignment are randomly and equally divided over all groups. We divided these non-responding students equally across the groups for two reasons. Practically, when the assignment was conducted, it was unsure whether students in the end choose to start an education program in this Law School. Those students who fill in the survey almost always attend a study program, while this is not the case for the non-responding students. Therefore spreading them equally avoids the issue of reallocating students afterwards because some groups would contain too many students. Secondly, since non-responding students may differ from responding students in non-observable ways in their peer effect, it is important to divide them equally across groups in order to avoid this bias.

There are two other restrictions regarding group assignment, which follow from university policy. First, all students within a work group need to have chosen the same study specialization. Second, between work groups of the same specialization the gender balance should be the same. Therefore, the randomization into the groups: high, mixed, and low creativity is based on the creativity type classification, study specialization, and gender of the student. Fig. 2 shows the final creativity distribution per group. Distributional tests show that the distribution clearly differs across each group type, be it low, mixed, and high creativity groups.[8]

In addition to differences in creativity between the low, mixed, and high creativity group, we also expect some differences in other characteristics, because individuals with different creativity scores show minor differences in other characteristics (see also Table 2). Therefore, in Table 3 we compare characteristics of the students assigned to the different group types. We find some small differences in characteristics of students assigned to the different creativity groups. Most notably, we see that in higher creativity groups students have higher grades in the Dutch language. In most of our regressions we control for all these different individual and group characteristics.

4.3. Non-participation

Since the aim of this paper is to study the impact of peer creativity on study performance, students only enter our main sample when they obtain a creativity score through survey participation. The participation rates range from 60% to 70% per cohort. We find some selection into survey participation, see Table A.1. Female students and students with better high school grades are somewhat more likely to participate in the survey, and there are some differences between study programs. As discussed before, the distribution of non-participating students across work groups is by design the same. Therefore, even if non-participating students differ in dimensions important for other students' productivity in work groups, this non-participation does not lead to a bias in our estimate of peer effects.

[8] We use the Kolmogorov–Smirnov test and compare each group with the other separately. We find in all three comparisons a p-value smaller than 0.01. This implies clear creativity differences between the three groups.

Table 3
Descriptive statistics: group types.

	L-group	M-group	H-group	p: L-M	p: L-H	p: M-H
Male (%)	36.38	35.12	30.80	0.684	0.048**	0.141
HS Grade Dutch language	6.72 (0.67)	6.74 (0.69)	6.84 (0.67)	0.715	0.007***	0.026**
HS Grade Mathematics	6.63 (1.06)	6.71 (1.15)	6.76 (0.933)	0.272	0.038**	0.473
Persistence	23.18 (2.74)	23.09 (2.55)	22.90 (2.80)	0.594	0.090*	0.257
Risk preference	6.94 (1.54)	7.03 (1.39)	6.87 (1.28)	0.365	0.408	0.061*
Observations	663	564	746			

The full sample refers to the sample of students who participated in the RAT (creativity test). L-group, M-group, and H-group refer to students that are assigned to a low, mixed, and high creativity group respectively. The standard deviations are in parentheses.

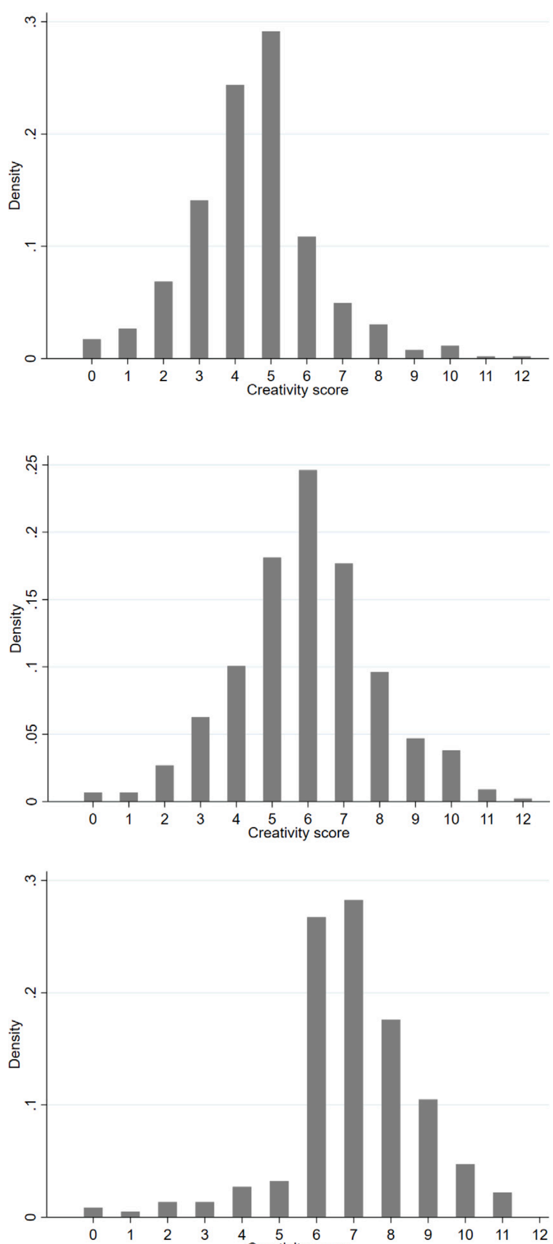

Fig. 2. The creativity distribution for low (top panel), mixed (middle panel), and high (bottom panel) creativity groups respectively. Creativity is measured on a scale from 0 to 12.

Table 4
Randomization test: group assignment on individual characteristics.

Dep. var:	HS Dutch	HS Math	Male	Risk	Persistence
A_i low	−0.052 (0.038)	−0.012 (0.036)	0.009 (0.040)	0.012 (0.164)	−0.117 (0.113)
A_i high	0.026 (0.032)	0.023 (0.029)	−0.038 (0.042)	−0.127 (0.154)	−0.128 (0.091)
Controls	Y	Y	Y	Y	Y
Observations	1973	1973	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type (high or low), and creativity score (continuous 0–12). A_i refers to the group type. A_i low means that the student is assigned to a low creativity group, and A_i high means that the student is a assigned to a high creativity group. The baseline is a mixed creativity group. The standard errors are clustered by work group and are in parentheses.

4.4. Randomization checks

In Table 2 we have seen some minor differences in the characteristics of high creativity students as compared to low creativity students. Particularly, high creativity students are more often female and have slightly better high school grades in the Dutch language. In addition – in Fig. 2 – we have seen clear differences in the creativity of peers in the low, mixed, and high creativity group. In this section we test whether the randomization has worked by conducting two tests. First, we test whether the assignment to the own type versus mixed type creativity group is unrelated to the individual characteristics conditional on creativity (Table 4). Second, we test whether the assignment to the own type versus mixed type creativity group is unrelated to peer group characteristics other than creativity (Table 5). If there would be a strong correlation between the randomly assigned creativity peer group type and individual and/or peer characteristics (while controlling for the creativity of the individual), than these other characteristics could drive the effect of a peer group on study achievement.

It is important to control for both the continuous as well as discrete creativity score, because they both impact the individual and peer group characteristics. The continuous creativity score correlates with for example the high school grades (the higher the creativity score the higher the grades), and the discrete score impacts the allocation to a peer group. The high creativity students can only be assigned to a high creativity or a mixed creativity group, while the low creativity students can only be assigned to a low creativity or a mixed creativity group. Therefore also the discrete score impacts the peer characteristics. Hence, since we want to learn whether the randomization has worked, we need to control for both the continuous and discrete creativity score. The results in Table 4 show (as expected) no differences in individual characteristics. This implies that the randomization has worked as intended.

We now check whether the randomization resulted in differences in peer characteristics other than creativity. If other peer characteristics differ between students in high, mixed, and low creativity groups,

Table 5
Randomization test: group assignment on peer characteristics.

Dep. var.: leave-out mean	HS Dutch	HS Math	Gender	Risk	Persistence
A_i low	−0.158	−0.001	0.184	−0.192	0.244
	(0.221)	(0.229)	(0.221)	(0.195)	(0.221)
A_i high	0.143	0.264	0.107	−0.194	−0.227
	(0.216)	(0.217)	(0.231)	(0.164)	(0.208)
Controls	Y	Y	Y	Y	Y
Observations	1973	1973	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type (high or low), and creativity score (continuous 0–12). A_i refers to the group type. A_i low means that the student is assigned to a low creativity group, and A_i high means that the student is a assigned to a high creativity group. The baseline is a mixed creativity group. The standard errors are clustered by work group and are in parentheses.

we may not be able to distinguish the impact of academic achievement that comes from peer creativity from the impact of other peer characteristics. For this we run an OLS regression where we regress the allocation to a high or low creativity peer group (with a mixed group as the reference category) on the leave-out mean of other peer characteristics: high school grades in Dutch and Math, gender, risk preferences, and persistence. We control for the continuous creativity score, and for the creativity type (high or low creativity score). Table 5 shows no significant differences in peer characteristics other than creativity. However, in line with the descriptive statistics in Tables 2 and 3 we see low creativity groups seem to have lower Dutch grades and high creativity groups have higher Dutch grades.[9]

5. Empirical methodology

In order to estimate the impact of peer creativity on students' academic performance we follow two distinct strategies. We first study the impact of the allocation to a creativity group type, and secondly, we study the impact of actual peer creativity on academic performance.

Students with a low creativity score are allocated to either a low creativity or mixed creativity peer group (where low and high creativity students are mixed), and students with a high creativity score are similarly allocated to either a high creativity or mixed creativity group. We therefore compare students who are allocated to their own creativity type to a mixed creativity type group. Specifically, we estimate:

$$Y_i = \delta A_i + \pi X_i + \kappa X_{-i} + \eta i \tag{1}$$

where Y_i is academic achievement measured in various ways: by grade point average (GPA), by number of courses completed, and by whether the student drops out of the first year, and Y_i is mean centered by cohort.[10] A_i is a dummy that equals 1 if the student is assigned to a group containing their own creativity type of students (low types for low creativity students and high types for high creativity students), and equals 0 if the student is assigned to a mixed creativity group. X_i is a vector of control variables and include: gender, study specialization, cohort, high school grades in Dutch and Math, risk preferences, and persistence. Finally, X_{-i} is a vector with the leave-out means of peer characteristics and consists of the gender composition, high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered at the work group level.[11]

Roughly 30% of the students fill in the survey only after the randomization. As a consequence they may be allocated to the 'wrong' type of peer group. For example, some high creativity students end up in low creativity groups and vice versa. In addition, just from the randomization we have that some low creativity groups contain a lower average creativity than other low creativity groups (and similarly for the mixed and high creativity groups there are creativity differences within the same group type). We can also make use of this variation. Therefore as a second strategy we study how the average of peers' creativity impacts study performance. To be precise, we estimate:

$$Y_i = \alpha C_i + \beta \overline{C_{-i}} + \gamma X_i + \xi X_{-i} + \epsilon_i \tag{2}$$

where Y_i, and X_{-i} are defined in the same way as in Eq. (1). The vector X_i contains (in addition to gender, high school grades, study specialization, cohort, risk preferences, and persistence) now also student i's creativity type. C_i is the creativity score of individual i and $\overline{C_{-i}}$ is the leave-out mean of creativity of student i's peer group (i.e. the mean of creativity of the group, without the creativity score of the individual). C_i and $\overline{C_{-i}}$ are both mean centered by cohort. Finally, the standard errors are clustered at the work group level.

6. Results

6.1. The impact of group allocation

In Table 6 we provide the estimates of A_i from Eq. (1). In the top panel we use the sample of low creativity students. These are the students who scored between 0 and 5 points on the RAT before randomization. A_i is a dummy that equals zero if the student is assigned to a mixed group (i.e. a group consisting of a mix of high and low creativity students) and A_i equals 1 if the student is assigned to a group with only low creativity students. The students who have not filled in the survey at the time of randomization are assigned equally to all groups and hence are not part of this sample. This means that A_i gives the effect of assigning a student to students who are similar in terms of creativity as compared to a mixed group. Column 1 shows the effect without any controls, in column 2 we control for individual characteristics, and in column 3 we also control for both individual and peer group characteristics. The top panel shows that for low creativity students there is no significant difference for those who are assigned to a low creativity peer group as compared to a mixed creativity group. However, the point estimates all suggest that if anything the effect of a lower creativity peer group is negative (as expected).

The bottom panel shows – for the sample of high creativity students – the impact of being assigned to a high creativity peer group as compared to a mixed creativity group. The results show sizable significant effects. Being assigned to a high creativity peer group leads to an increase in GPA of 8.1 – 11.5 percentage points.[12] The reason that the positive effect of creative peers on study performance is stronger for high creativity students may be that one needs a certain creative ability to understand creative questions and use creative approaches of work group members. The results are qualitatively similar if we take the number of courses or whether students drop-out of the program as outcome variables, see Tables A.2 and A.3.

[9] If we run the same regression as reported in Table 5 with the leave-out mean of peer creativity we obtain sizable and strongly significant coefficients,to be precise for the low group: −0.606*** (0.194), and for the high group: 0.467*** (0.156). This confirms the graphical evidence in Table 2, that the randomization procedure let to sizable differences in the creativity of peers.

[10] When we measure GPA we treat missing grades as a 1.0 grade. The reason for this is that this is the actual grade that students obtain when they do not fill in anything in an exam.

[11] The full sample consists of 159 work groups.

[12] These results are robust to using randomization inference.

Table 6
The effect of group allocation on GPA.

Dep. var.: GPA	L-type	L-type	L-type
A_i	−0.067	−0.058	−0.069
	(0.057)	(0.057)	(0.056)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	586	586	586
Dep. var.: GPA	H-type	H-type	H-type
A_i	0.115*	0.088*	0.081*
	(0.061)	(0.044)	(0.042)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	807	807	807

This table shows the results of an OLS regression where the individual controls are students' creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

6.2. The impact of peer creativity

Since close to 30% of the students participate in the survey only after the randomization was completed (but before starting at University, and before meeting any students) some students end up in the 'wrong' group type. In other words, some high creativity students end up in a low creativity group, and vice versa. Therefore we next use Eq. (2) to estimate the impact of the mean of peer creativity on students' academic outcomes.

We build up Eq. (2) in Table 7, the first column shows the impact of creativity and peer creativity without controls, in column 2 we add individual control variables, and in column 3 we also add peer group controls. The results show a positive correlation between an individual's creativity (while controlling for the creativity of the peer group) and study performance. When we control for individual and peer control variables we find that a 1 standard deviation increase in creativity is associated with a 8.3 percentage points increase in GPA. This effect is smaller than the correlation between creativity and GPA which equals 0.136.

The variable of interest is (the leave-out mean of) peer creativity. We find that a 1 standard deviation increase in the mean of peer creativity leads to an increase of at most 11.4 percentage points (without using any control variables) to 7.6 percentage points when we control for individual characteristics (specifically: gender, study specialization, cohort, and high school grades), and finally 6.2 percentage points when we control for both individual and peer group controls (specifically the leave-out-mean of gender, high school grades, risk and persistence). If we compare the first and second column, we see that the effect of peer creativity is somewhat smaller when we include individual control variables. The reason is that students with better high school grades (and women) have a higher GPA, but also score slightly higher on the creativity test (and are because of their higher creativity score assigned to more creative peers). When we include also peer group control variables, the effect of more creative peers decreases only slightly. This is evidence that it is not mostly language (or math) expertise of peers, or risk preference or persistence of peers, that leads to the increase of GPA, but the creativity of peers.

Our results are compared to the literature somewhat smaller than the impact of peer persistence in Zou (2024), but larger than the impact of risk preference and persistence in Golsteyn et al. (2021). The results are qualitatively similar if we take the number of courses or whether students drop-out of the program as outcome variables, see Tables A.4 and A.5.

Table 7
The effect of peer creativity on GPA.

Dep. var.: GPA	(1)	(2)	(3)
C_i	0.036	0.083*	0.083*
	(0.040)	(0.045)	(0.045)
$\overline{C_{-i}}$	0.114***	0.076**	0.062**
	(0.029)	(0.031)	(0.032)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

6.3. Heterogeneous effects

We now conduct some heterogeneity analysis, which can be considered an exploratory analysis.[13] We study whether the higher creativity students benefit more from creative peers (as the analysis in the previous subsection suggested), and whether students with higher (or lower) academic ability benefit, which we study by comparing students above and below median high school grades.

Table A.6 shows – in line with the previous subsection – that the positive creativity peer effects for the high creativity type of students are strongly significant and relatively large (varying from 14.4% to 7.8% of a standard deviation increase in GPA following a 1 standard deviation increase in peer creativity. For the low creativity type students the effects are much weaker (8.6%–2.9%) and not always significant. This robust finding (that mostly high creativity students benefit from creative peers) is in line with the idea that students need to be creative themselves in order to benefit from the creative ideas and approaches from their peers. Students who lack creativity may not understand how a creative approach of a peer can be used in a slightly different context. As an alternative approach we re-estimate Eq. (2) but with the inclusion of an interaction term between the individual creativity and leave-out-mean creativity. We then estimate from that equation the marginal effect of a 1 standard deviation increase in peer creativity for different levels of an individual's creativity. In Fig. A.2 we plot the coefficients. This shows – in line with the previous results – that the higher creativity students benefit (most) from having more creative peers.

Table A.7 shows the results split up by high school grades. These results show that the students with a higher academic ability at baseline benefit more from creative peers.

Table A.8 shows the results split up by cohort. We find no significant difference between the years, even though cohort 20–21 spent a large fraction of their work groups online (because of COVID).

6.4. Robustness: Correlates of Creativity

So far we have shown that students who are randomly assigned to more creative work group peers perform better at university. Creativity may be correlated with cognitive and/or non-cognitive skills that result in peer effects. It could therefore be the case that peer creativity impacts study performance through a trait or skill that is correlated with creativity, instead of creativity itself. In some of our regressions we control for peers' high school grades (see Booij et al. (2017) for evidence of peer effects of high school study grades on academic achievement at university), and for risk preferences and persistence (see Golsteyn et al. 2020 for evidence of peer effects of these preferences/skills). Our main

[13] While the earlier analysis is in line with the pre-analysis plan, this heterogeneity analysis is added later to the paper. See also the final section of Appendix C.

Table 8
The effect of peers' stated creativity on GPA.

Dep. var.: GPA	(1)	(2)	(3)
C_i	0.004	0.003	0.001
	(0.020)	(0.021)	(0.021)
$\overline{C_{-i}}$	0.087**	0.024	0.011
	(0.028)	(0.026)	(0.027)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are gender, high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

result: the peer effects of creativity on study performance survive when we add all these other peer characteristics, see column 3 of both Tables 6 and 7.

Other cognitive and non-cognitive skills that are known to be correlated with creativity are intelligence (see e.g. Lee and Therriault (2013)) and the big five personality traits (mainly openness to experience), see e.g. Pesout and Nietfeld (2021). We do not have measures of these skills for our subjects. However, we take a 'convenience sample' of 95 students – who were willing to participate in a survey – from the same university programs that started in the academic year 2022–2023 for which we measure IQ, the big five personality traits, and creativity using the RAT. For this sample we establish three facts. First, there is only a weak and insignificant correlation between the big five personality traits and creativity, where openness to experience is marginally significant, and the other traits are not. Second, there is a significant correlation between IQ and creativity. Third, this positive correlation disappears once we control for high school grades. In the baseline specification we control for high school grades and for high school grades of peers. Therefore we believe there are no strong correlates with creativity that likely drive the creativity peer effect that we find. Taken together our findings mirror the statement in Gill and Prowse (2024) that creativity is a unique trait that affects educational attainment besides the standard cognitive and non-cognitive traits.

6.5. Stated creativity

Research has found that the correlation between stated creativity and academic achievement is positive but weaker than the correlation between revealed creativity and academic achievement, see e.g. Gajda et al. (2017). However, when studying peer effects of creativity it is crucial that students show their creativity, e.g. through asking questions, participating in class, or in group assignments. Hence, some creative students – as measured by the RAT – may not believe they are creative and hence contribute less to discussions leading to less positive spillovers to their peers, this may lead to a stronger correlation between peer creativity measured by a stated creativity test and academic performance. On the other hand, some students may be overconfident and believe they are creative, while in fact they are not. This may lead to a weaker correlation between peer effects of stated creativity and academic performance.

We study the peer effects of a stated creativity measure, the K-DOCS which we adjust to fit the university context, see Section 3.2 for more the questions and their scales. We estimate the impact that the stated creativity of peers has on students using Eq. (2). We mean standardize the stated creativity variable, in the same way as the revealed creativity measure. Note that the randomization is based on the RAT, not on the K-DOCS, hence the between group variation in K-DOCS creativity is less pronounced. Fig. A.1 shows the distribution of the domain specific K-DOCS scores across the different type of work groups.

Table 8 shows that the impact of stated creativity on students' own study performance is small, positive, and insignificant. The peer effects are positive, mostly insignificant, and smaller as compared to the RAT measure of creativity. The fact that our results using the alternative K-DOCS measure are weaker than with the RAT measure is in line with the practice that students are assigned to a work group based on their RAT score not based on their K-DOCS score, and is in line with the literature that shows a weaker correlation between self-stated creativity measures and study performance. Another reason for the difference between the RAT and K-DOCS results may be related to overconfidence. We see that women (as compared to men) have both a higher RAT score and better study grades, but a lower K-DOCS score. Given that the literature (see e.g. Adamecz-Völgyi and Shure (2022)) has shown that men are more overconfident than women, the higher K-DOCS score of men may come from overconfidence instead of creativity.

6.6. Mechanisms

So far we have established that there are positive peer effects from creativity on students' academic performance. In this section we try to distinguish between some of the mechanisms that may drive these results. We study three mechanisms: changes effort (i.e. in number of study hours), changes in hourly productivity, and a change in creative ability.

Hourly productivity may increase because the novel questions and discussion improve students' understanding during class and may teach them learning strategies they can use outside of class to study more efficient. Hourly productivity may also decrease when (especially less creative) student do not understand the creative questions and discussions in class. This may decrease the understanding of the study material without adjusting effort.

Effort may increase because of increased motivation and engagement of students who are motivated by the novel questions and discussions during class. Alternatively, effort may also decrease if students become confused and discouraged from discussions that they find hard to understand.

Finally, students creative abilities may change as a consequence of exposure to creative peers, this may improve their study performance. This last hypothesis – that creativity improves after exposure to creative student – is in line with Shan and Zölitz (2022) who find peer effects in the big five personality traits. Students with peers who are more conscientious and have a higher openness to experience improve these traits over time, but not the other traits. Shan and Zolitz argue that students adopt the productive traits, but not the non-productive traits.

In the second semester, roughly six months after the start of the study program (see also Fig. 1) we administered a survey containing questions that help us distinguish between the three mechanisms described above. We ask about work group dynamics in order to learn whether students believe groups with more creative peers are more productive. In order to measure these group dynamics we ask respondents to rate the following two statements on a five-point-scale, ranging from 1 = never to 5 = always: *I think I learn a lot from studying together with students from my work group*, and *I spend time with students from my work group outside university*. If students experience that creative peers contribute to work groups in a way that their productivity per hour increases, we would expect them to respond more positive to the statement about whether they learn a lot from students in their work group. In order to distinguish whether increased productivity comes from an increase in productivity per hour or from increased effort, we also ask students about the amount of hours that they study, both with their work group peers as well as the total number of hours. Finally, the second survey contains another RAT (but with different questions) and another K-DOCS. From the RAT we obtain a new measure of creativity that we can use to analyze whether students with more creative peers become more creative. We analyze whether the mean peer creativity (as measured by the RAT) impacts the student's survey responses

Table 9
Selection into follow-up survey.

Dep.var. participation in survey 2			
A_i	−0.014	0.017	−0.022
	(0.031)	(0.045)	(0.040)
Creativity type	0.000		
	(0.053)		
C_i	−0.006	0.084**	−0.062**
	(0.026)	(0.040)	(0.031)
$\overline{C_{-i}}$	0.007	0.030	−0.006
	(0.018)	(0.032)	(0.023)
Controls	Y	Y	Y
Observations	1393	586	807

This table shows the results of an OLS regression where the individual controls are gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are gender, high school grades in Dutch and Math, risk preferences, and persistence. Column 1 consists of the full sample of students who filled in the survey before randomization, column 2 and 3 consist of the sample of low creativity and high creativity type students respectively. A_i is the assignment to a creativity group type. In column 1 A_i is continuous (with low = −1, mixed = 0, high = 1), in column 2 and 3 A_i is the low and high creativity group respectively (with the mixed group as the baseline). The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

while also controlling for initial creativity as well as creativity type, using OLS regressions. We use the K-DOCS to see whether exposure to more creative peers (also) increases students' self-concept of creativity (i.e. their beliefs about their own creativity).

Roughly 20% of the students who participated in the baseline survey also participated in the follow-up survey. Table A.9 shows that students who are more able (at baseline and in performance at university) are more likely to participate in the second survey. Particularly, students who participate in the first and second survey are: more likely to be female, have better high school grades, and have at university a higher GPA, pass more courses, and are less likely to drop-out of the study program. However, to test for mechanisms, it is important that peer creativity is randomized. We therefore test whether students from certain peer groups are more (or less) likely to participate in the second survey. If only or mainly those assigned to a certain work group participate in the survey our results would be biased. In column 1 of Table 9 (the full sample) we find no indication of selection into survey participation based on the allocation of students into groups. Neither the allocation to a creativity group type, nor peer creativity measured before randomization are significantly correlated with the decision to participate in the second survey. Columns 2 and 3 with only the low and high creativity type shows no significant impact of the allocation to a creativity group type or average peer creativity.

Table 10 shows how (the mean of) peer creativity relates to students' responses to the second survey. In column 1 we regress (peer) creativity on students' response to the questions whether they 'believe they learn a lot from other students in their work group. We find a small and insignificant (but positive) effect. Similarly, in column 2 we see that peer creativity appears unrelated to students' response to the question whether they spent time with work group members outside the classroom. These responses suggest that students are unaware of the positive impact that creative peers have on their study performance. Columns 3 and 4 show the impact of peer creativity on the time that students spent studying. These results show that there is no impact on the total number of hours students study, nor on the allocation of these study hours. This suggests that students do not adjust their effort (measured in study hours) based on the creativity of their peers. In column 5 we see that students with one standard deviation higher creativity peers, score 26.6 percentage points higher on the RAT in the second survey. This responds to around half an extra question correct on the RAT on average. This suggests that students who are exposed to high creativity peers become more creative.[14] The finding that students' environment (i.e. the interaction with their peers) can affect their creativity is in line with Tan, Chin, Chng, Lee, and Ooi (2022) who find that students' perceived social support can boost creativity. Finally, in column 6 we study the impact of peer creativity and students' self-concept of creativity. The idea is that students who become more creative (measured by the RAT) may also believe they have become more creative. Perhaps surprisingly, we do not find this effect. We see no significant change in the K-DOCS score. This result is in line with the fact that the correlation between the RAT and K-DOCS score is also small.[15] If we split the sample by creativity type, we find similar results for both the low and high creativity student type, see Tables A.10 and A.11.

7. Discussion and conclusion

In this paper we studied the impact of peer creativity on academic achievement. We find that the distribution of peer creativity is important for study performance. Students with a one standard deviation more creative peer group obtain a 6.2 to 7.4 percentage point higher GPA. We subsequently explore what drives these creativity peer effects. Our findings suggest that students do not increase the number of hours they study, and hence become more productive per hour. Surprisingly, students seem to be naive about these positive effects given that they do not believe they learn more from their creative peer group. In addition, creative peers seem to help students improve their creative ability half a year later.

This research has several implications. First, the fact that peer creativity positively impacts other students' academic achievement implies that creativity can be part of selection criteria into selective universities. This holds especially since creativity is only weakly correlated with the variables that schools mostly use for selection such as prior study grades. Second, the fact that there are positive spillovers from creative students to their peers (i.e. students become more creative from interactions with creative peers), shows that there are benefits of creative peers that (likely) exceed the education phase and have longer run labor market implications. These arguments strengthens the case to have creativity training as part of an academic curriculum. See for instance Fleith, Renzulli, and Westberg (2002), Morin, Robert, and Gabora (2018), and Ritter, Gu, Crijns, and Biekens (2020) for suggestions about creativity training programs, and their impact on students' creative abilities. Third, since students seem unaware of the advantages that studying with more creative peers has, it is useful to reveal this information to students. This may trigger students to reallocate effort towards studying together with creative peers which increases the positive peer effects. Fourth, our findings have implications for the allocations of students to smaller work groups. We find an indication that everyone benefits from creative peers, but that most of the gains are for students who themselves are of high creativity, and those with better high school grades (i.e. prior academic ability). If universities aim to improve the performance of high creativity students or the average performance of all students, they should choose to create homogeneous groups (i.e. create high and low creativity work groups). If, however, the goal is to boost performance of the low creativity and low high school grades students, then creating mixed groups may be better.

While this paper finds robust peer effects of creativity and can shed some light on the reasons for the increased performance of students with more creative peers, additional research is desired. First, different creativity measures (such as divergent thinking tests) may

[14] We see that students score on average also slightly higher on the RAT. The mean score increases from 5.82 to 6.08.

[15] If we compare the K-DOCS scores from the first and second survey we see little change, the score changed from 29.58 tot 29.48.

Table 10
Exploration of mechanisms.

Dep. var:	(1)	(2)	(3)	(4)	(5)	(6)
C_i	−0.054	−0.174	−1.154	−1.236*	0.520**	0.136
	(0.182)	(0.159)	(2.043)	(0.698)	(0.205)	(0.120)
$\overline{C_{-i}}$	0.126	0.016	−2.169	−0.467	0.266**	0.021
	(0.108)	(0.096)	(1.450)	(0.474)	(0.117)	(0.098)
Individual controls	Y	Y	Y	Y	Y	Y
Peer controls	Y	Y	Y	Y	Y	Y
Observations	384	384	384	384	384	384

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively. The dependent variable is: (1) whether the student believes he learns a lot from other work group members, (2) whether the student spends time outside university with work group members, (3) the number of hours studied‘with workgroup student, and (4) the total number of hours studied, and (5) creativity measured in the follow up survey with the RAT, (6) creativity measured in the follow up survey with the K-DOCS.

Table A.1
Descriptive statistics: response versus non-response first survey.

	Respondent	Non-respondents	p-value
Male (%)	35.48	46.36	0.000
HS Grade Dutch language	6.72	6.62	0.000
	(0.68)	(0.67)	
HS Grade Mathematics	6.65	6.56	0.015
	(1.02)	(0.99)	
GPA	5.15	4.65	0.000
	(1.96)	(2.09)	
Course credits	5.46	4.68	0.000
	(3.55)	(3.59)	
Drop out (%)	20.1	23.0	0.029
Observations	1973	1709	

The sample of respondents refers to the full sample used for analyses. The non-respondents are those students who did not complete the survey.

Table A.2
The effect of group allocation on course completion.

Dep. var.: nr. of courses	L-type	L-type	L-type
A_i	−0.373	−0.398*	−0.466*
	(0.274)	(0.233)	(0.238)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	586	586	586
Dep. var.: nr. of courses	H-type	H-type	H-type
A_i	0.238	0.132	0.055*
	(0.290)	(0.204)	(0.191)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	807	807	807

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.3
The effect of group allocation on drop out.

Dep. var.: drop out	L-type	L-type	L-type
A_i	0.002	0.002	0.003
	(0.002)	(0.002)	(0.003)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	586	586	586
Dep. var.: drop out	H-type	H-type	H-type
A_i	−0.004	−0.003	−0.004
	(0.004)	(0.004)	(0.004)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	807	807	807

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

strengthen the robustness of our creativity measure for our results. Second, observed study behavior from test preparations and exam taking (in addition to stated behavior from the survey) would give more reliability to the claim that students become more productive per hour. Third, while our peer effects are based on the entire work group, students generally do not interact with all students to the same extent. Information on students' social network and/or randomization of students' desk assignment (such as in Harmon, Fisman, and Kamenica (2019), Wu, Zhang, and Wang (2023)) within class may provide additional insights in what drives the peer effects. Fourth, studying the behavior of students in class (for instance through video recordings) can help us understand what it is that leads students with creative peers to become more creative. For example, is it discussions between students or are students learning from questions that creative students ask to their teacher. Finally, it is interesting to learn to what extent our findings generalize to other settings. For instance, in our context students participate in the same work group for the entire first year, hence there is sufficient time to learn from each other. Future research is needed to learn whether shorter interactions – for instance in cases where students switch between work groups per course – can have similar effects. In addition, it is interesting to learn to what extent our findings generalize to a labor market setting where people produce in teams.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Tables

See Tables A.1–A.11

Table A.4
The effect of peer creativity on course completion.

Dep. var.: nr. of courses	(1)	(2)	(3)
C_i	0.033	0.143	0.151
	(0.117)	(0.165)	(0.166)
$\overline{C_{-i}}$	0.539***	0.311**	0.251**
	(0.133)	(0.123)	(0.119)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.5
The effect of peer creativity on drop out.

Dep. var.: drop out	(1)	(2)	(3)
C_i	−0.008	−0.001	−0.003
	(0.012)	(0.018)	(0.019)
$\overline{C_{-i}}$	−0.028***	−0.011	−0.008
	(0.011)	(0.013)	(0.013)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1973	1973	1973

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.6
The effect of peer creativity on GPA.

Dep. var.: GPA	(1)	(2)	(3)
Top panel: Low creativity type students			
C_i	0.045	0.058	0.051
	(0.67)	(0.071)	(0.071)
$\overline{C_{-i}}$	0.086*	0.037	0.029
	(0.45)	(0.051)	(0.054)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	831	831	831
Bottom panel: High creativity type students			
C_i	0.130**	0.091*	0.099*
	(0.057)	(0.054)	(0.054)
$\overline{C_{-i}}$	0.144***	0.094**	0.078**
	(0.036)	(0.037)	(0.038)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1142	1142	1142

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.7
The effect of peer creativity on GPA.

Dep. var.: GPA	(1)	(2)	(3)
Top panel: Low HS grade students			
C_i	−0.018	0.038	0.038
	(0.044)	(0.071)	(0.071)
$\overline{C_{-i}}$	0.032	−0.013	−0.011
	(0.039)	(0.044)	(0.045)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	832	832	832
Bottom panel: High HS grade students			
C_i	0.074*	0.136**	0.139**
	(0.039)	(0.055)	(0.056)
$\overline{C_{-i}}$	0.159***	0.128***	0.105***
	(0.036)	(0.037)	(0.039)
Individual Controls	N	Y	Y
Peer controls	N	N	Y
Observations	1141	1141	1141

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.8
The effect of peer creativity on GPA.

Dep. var.: GPA	(1)	(2)	(3)
Cohort:	19–20	20–21	21-22
C_i	0.091	0.121	0.023
	(0.087)	(0.083)	(0.079)
$\overline{C_{-i}}$	0.042	0.051	0.082
	(0.069)	(0.055)	(0.060)
Individual Controls	Y	Y	Y
Peer controls	Y	Y	Y
Observations	622	694	657

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively.

Table A.9
Descriptive statistics: response versus non-response second survey.

	Respondent	Non-respondents	p-value
Male (%)	26.56	37.63	0.000
HS Grade Dutch language	6.88	6.68	0.000
	(0.64)	(0.68)	
HS Grade Mathematics	6.85	6.60	0.000
	(1.18)	(0.97)	
GPA	6.37	4.86	0.000
	(1.17)	(1.99)	
Course credits	7.59	4.94	0.000
	(2.61)	(3.56)	
Drop out (%)	0.4	25.0	0.000
Observations	384	1589	

The sample of respondents refers to the students who filled in the second survey. The non-respondents are those students who did not complete the second survey.

Table A.10
Exploration of mechanisms for low creativity students.

Dep. var:	(1)	(2)	(3)	(4)	(5)	(6)
C_i	0.068	−0.248	−0.215	−1.753	0.552	0.502
	(0.347)	(0.302)	(2.409)	(1.265)	(0.357)	(0.120)
$\overline{C_{-i}}$	0.120	0.092	−0.495	0.536	0.324*	0.062
	(0.191)	(0.176)	(1.782)	(0.870)	(0.178)	(0.153)
Individual controls	Y	Y	Y	Y	Y	Y
Peer controls	Y	Y	Y	Y	Y	Y
Observations	155	155	155	155	155	155

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively. The dependent variable is: (1) whether the student believes he learns a lot from other work group members, (2) whether the student spends time outside university with work group members, (3) the number of hours studied'with workgroup student, and (4) the total number of hours studied, and (5) creativity measured in the follow up survey with the RAT, (6) creativity measured in the follow up survey with the K-DOCS.

Table A.11
Exploration of mechanisms for high creativity students.

Dep. var:	(1)	(2)	(3)	(4)	(5)	(6)
C_i	−0.115	−0.067	−0.610	−0.885	0.609***	0.172
	(0.221)	(0.196)	(2.267)	(0.882)	(0.223)	(0.257)
$\overline{C_{-i}}$	0.128	0.001	−3.437	−0.779	0.336**	0.168
	(0.141)	(0.129)	(2.239)	(0.542)	(0.160)	(0.120)
Individual controls	Y	Y	Y	Y	Y	Y
Peer controls	Y	Y	Y	Y	Y	Y
Observations	215	215	215	215	215	215

This table shows the results of an OLS regression where the individual controls are creativity type, gender, study specialization, cohort, risk preferences, persistence, and high school study grades in Dutch and Math. The peer controls are peers' high school grades in Dutch and Math, risk preferences, and persistence. The standard errors are clustered by work group and are in parentheses. *, **, *** indicate significance levels at the 0.10, 0.05, and 0.01 level respectively. The dependent variable is: (1) whether the student believes he learns a lot from other work group members, (2) whether the student spends time outside university with work group members, (3) the number of hours studied'with workgroup student, and (4) the total number of hours studied, and (5) creativity measured in the follow up survey with the RAT, (6) creativity measured in the follow up survey with the K-DOCS.

Appendix B. Figures

See Figs. A.1 and A.2.

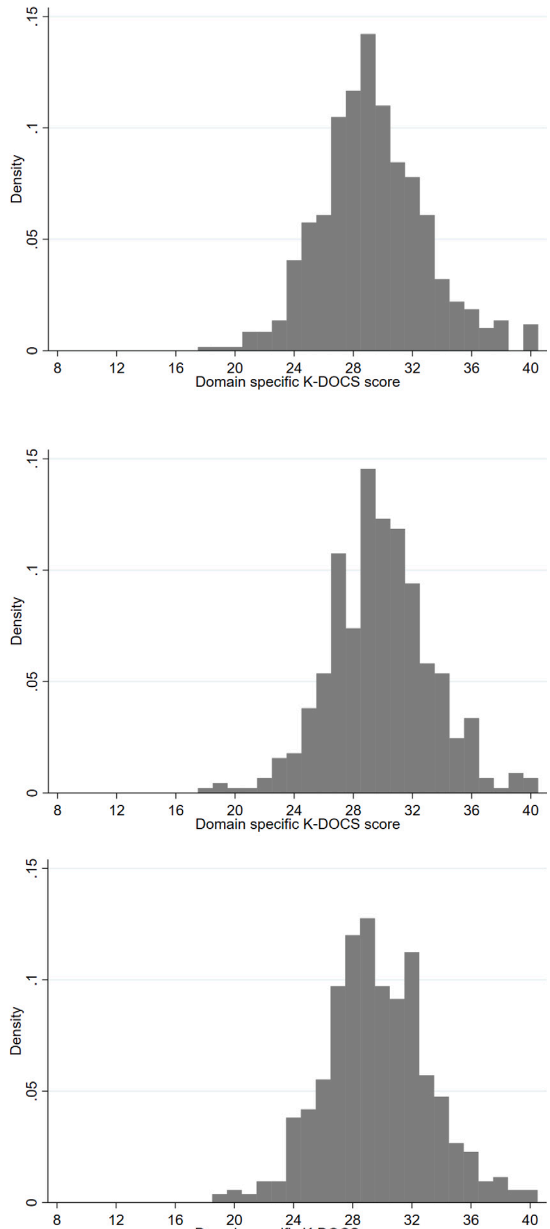

Fig. A.1. Stated creativity distribution for low, mixed, and high creativity groups respectively.

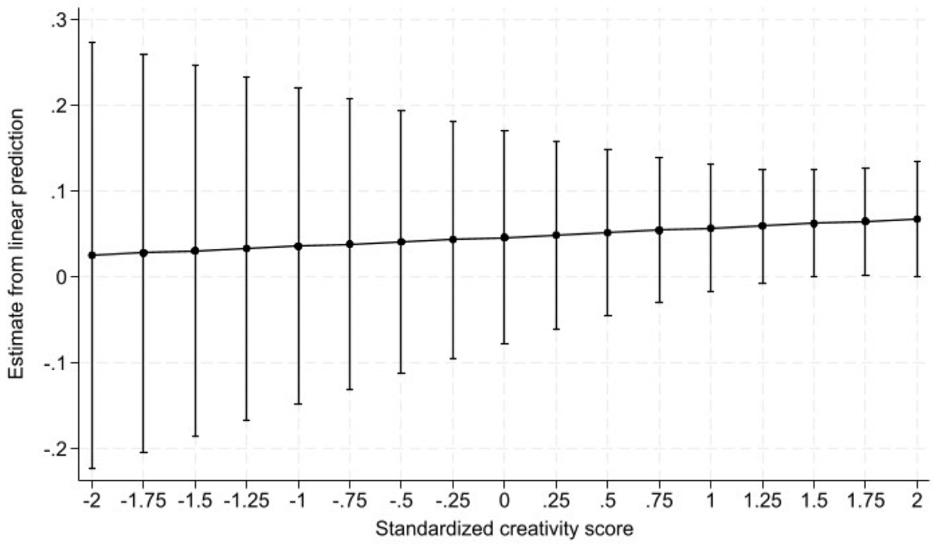

Fig. A.2. The average marginal effect of peer creativity.

Appendix C. Data collection, availability, and analysis

Data collection:

We use data from two sources: administrative data provided by the university, and data directly obtained from surveys (via the university). Since this paper is about the impact of creativity, students can only be part of this study when they obtain a creativity score through participation in a survey before the start of the academic year. Therefore what we refer to as the 'full sample' in the paper is the sample of students who participated in the first survey. Before the start of this study we obtained IRB approval. IRB approval was granted for the randomization of students into groups as well as for the analysis of students' administrative and survey data. The years before this experiment the randomization of students into work groups has been random (conditional on work groups specialization and an equal gender mix between groups). Learning from the outcomes of this experiment where we randomized students into work groups based on their creativity score, was considered more important than possible disadvantages such as some students ending up in mostly less creative groups, the additional effort students spent filling in survey questions, and a researcher working with administrative student data. Regarding the sensitivity of the data (in terms of students' privacy) we took two measures. First, the data remains stored at university and cannot be shared with other researchers. Second, students had the opportunity to ask questions about the use of their data, and could demand to be excluded from the study. This was communicated to the students when they received the survey invitation. No student made use of this opportunity. Data was collected on a yearly basis. Access to survey data was granted directly after the students filled in the data (see Fig. 1 in the paper), administrative data was made accessible as soon as possible after the end of the academic year (this was October or November in each year).

Deviations from the pre-analysis plan

Only a very limited pre-analysis plan was written. This included Eq. (2) in the final paper, including the use of the individual and peer control variables described in Section 5. In addition, Eq. (1): the impact of the assignment to creativity group types, is what directly follows from the experimental set up. The Section 6.3 (heterogeneous effects) is not analysis that was explicitly pre-registered, this can be considered exploratory analysis which followed after paper presentations and journal reviewer requests.

Appendix D. Remote Associates Test

The Remote Associates Test (RAT) is used as a way to measure creativity. It is a contingent thinking and verbal test. participants see three (seemingly unrelated) words. They are than asked to come up with a word associated to all three of the given words. The website: https://www.remote-associates-test.com/ provides a large data bank including the 3 words, the solution, and also an indication of the question's difficulty, ranging from very easy to very hard.

In our survey we first select 3 practice questions that students use in order to familiarize themselves with the problems, and they will see the correct answer to each questions after 20 s. Then the actual test follows in which students take 12 questions (1/3 of these questions are easy, 1/3 moderate, and 1/3 hard). Respondents receive 1 point per question for providing the correct answer (and 0 points for a wrong answer), and therefore can receive 12 points in total. Respondents have 20 s per question to come up with and fill in their answer, they also practice with this using the practice questions.

Examples of questions are: cottage/swiss / cake (with the solution: cheese), and worm/shelf / end (with the solution: book).

Data availability

All data used in the manuscript is privacy sensitive, and I agreed with the university (who owns the data) that these data will not be shared. Requests to run certain code on the data, for instance for replication purposes, can be send to the author (Max van Lent: m.van.lent@law.leidenuniv.nl).

References

Adamecz-Völgyi, A., & Shure, N. (2022). The gender gap in top jobs–the role of overconfidence. *Labour Economics*, *79*, Article 102283.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal*, *39*, 1154–1184.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98–108.

Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's metco program. *American Economic Review*, *94*(5), 1613–1634.

Atwood, S. A., & Pretz, J. E. (2016). Creativity as a factor in persistence and academic achievement of engineering undergraduates. *Journal of Engineering Education*, *105*(4), 540–559.

Backman, M. E., & Tuckman, B. W. (1972). Review of remote associates test. *Journal of Educational Measurement*, *9*(2), 161–162.

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

Bradler, C., Neckermann, S., & Warnke, A. J. (2019). Incentivizing creativity: A large-scale experiment with performance bonuses and gifts. *Journal of Labor Economics*, *37*(3), 793–851.

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

Carrell, S. E., 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.

CBS (2023). Al 23 jaar op rij meer vrouwen dan mannen in hoger onderwijs.

Charness, G., & Grieco, D. (2019). Creativity and incentives. *Journal of the European Economic Association*, *17*(2), 454–496.

Chermahini, S. A., Hickendorff, M., & Hommel, B. (2012). Development and validity of a dutch version of the remote associates task: An item-response theory approach. *Thinking Skills and Creativity*, *7*(3), 177–186.

De Giorgi, G., & Pellizzari, M. (2013). Understanding social interactions: Evidence from the class room. *The Economic Journal*, *124*(579), 917–953.

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), 1739–1774.

Erat, S., & Gneezy, U. (2016). Incentives for creativity. *Experimental Economics*, *19*(2), 269–280.

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.

Fleith, D. D. S., Renzulli, J. S., & Westberg, K. L. (2002). Effects of a creativity training program on divergent thinking abilities and self-concept in monolingual and bilingual classrooms. *Creativity Research Journal*, *14*(3–4), 373–386.

Freund, P. A., & Holling, H. (2008). Creativity in the classroom: A multilevel analysis investigating the impact of creativity and reasoning ability on GPA. *Creativity Research Journal*, *20*, 309–318.

Gajda, A., Karwowski, M., & Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*, *109*(2), 269.

Gill, David, & Prowse, Victoria (2024). The creativity premium: Exploring the link between childhood creativity and life outcomes. *Journal of Political Economy Microeconomics*, *2*(3), 495–526.

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

Gong, J., Lu, Y., & Song, H. (2021). Gender peer effects on students' academic and noncognitive outcomes: Evidence and mechanisms. *Journal of Human Resources*, *56*(3), 686–710.

Gross, D. P. (2020). Creativity under fire: The effects of competition on creative production. *The Review of Economics and Statistics*, *102*(3), 583–599.

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

Harmon, N., Fisman, R., & Kamenica, E. (2019). Peer effects in legislative voting. *American Economic Journal: Applied Economics*, *11*(4), 156–180.

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

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

Kaufman, J. C. (2010). Using creativity to reduce ethnic bias in college admissions. *Review of General Psychology*, *14*(3), 189–203.

Kaufman, J. C. (2012). Counting the muses: development of the kaufman domains of creativity scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 298.

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York, NY: Wiley.

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

Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence? Examining internal and external structure validity evidence of the remote associates test. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(4), 446.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*(5), 306–320.

Lyle, D. S. (2009). The effects of peer group heterogeneity on the production of human capital at west point. *American Economic Journal: Applied Economics*, *1*(4), 69–84.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220.

Mednick, M. T. (1963). Research creativity in psychology graduate students. *Journal of Consulting Psychology*, *27*(3), 265.

Morin, S., Robert, J. M., & Gabora, L. (2018). How to train future engineers to be more creative? An educative experience. *Thinking Skills and Creativity*, *28*, 150–166.

Mourgues, C., Tan, M., Hein, S., Elliott, J. G., & Grigorenko, E. L. (2016). Using creativity to predict future academic performance: An application of Aurora's five subtests for creativity. *Learning and Individual Differences*, *51*, 378–386.

Naderi, H., Abdullah, R., Aizan, H. T., Sharir, J., & Kumar, V. (2009). Creativity, age and gender as predictors of academic achievement among undergraduate students. *The Journal of American Science*, *5*(5), 101–112.

Naderi, H., Abdullah, R., Aizan, H. T., Sharir, J., & Kumar, V. (2010). Relationship between creativity and academic achievement: A study of gender differences. *The Journal of American Science*, *181*.

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

Pesout, O., & Nietfeld, J. L. (2021). How creative am i?: Examining judgments and predictors of creative performance. *Thinking Skills and Creativity*, *40*, Article 100836.

Ritter, S. M., Gu, X., Crijns, M., & Biekens, P. (2020). Fostering students' creative thinking skills by means of a one-year creativity training program. *PloS One*, *15*(3), Article e0229773.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far?. In *Handbook of the economics of education*: *Vol. 3*, (pp. 249–277). Elsevier.

Shan, X., & Zölitz, U. (2022). Peers affect personality development. *IZA Discussion Paper, No. 15257*.

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

Tan, C. S., Chin, X. Y., Chng, S. T. C., Lee, J., & Ooi, C. S. (2022). Perceived social support increases creativity: experimental evidence. *International Journal of Environmental Research and Public Health*, *19*(18), 11841.

Torrance, E. P. (1974). *The torrance tests of creative thinking-norms-technical manual research edition-verbal tests, forms a and b- figural tests, forms a and b*. Princeton, NJ: Personnel Press.

Vandevelde, S., Van Keer, H., & Rosseel, Y. (2013). Measuring the complexity of upper primary school children's self-regulated learning: a multi-component approach. *Contemporary Educational Psychology*, *38*, 407–425.

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

Wu, J., Zhang, J., & Wang, C. (2023). Student performance, peer effects, and friend networks: Evidence from a randomized peer intervention. *American Economic Journal: Economic Policy*, *15*(1), 510–542.

Yang, J., & Zhao, X. (2021). The effect of creative thinking on academic performance: Mechanisms, heterogeneity, and implication. *Thinking Skills and Creativity*, *40*, Article 100831.

Zárate, R. A. (2023). Uncovering peer effects in social and academic skills. *American Economic Journal: Applied Economics*.

Zhang, W., Ren, P., & Deng, L. (2020). Gender differences in the creativity–academic achievement relationship: A study from China. *The Journal of Creative Behavior*, *54*(3), 725–732.

Zimmerman David, J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *The Review of Economics and Statistics*, *85*(1), 9–23.

Zou, J. (2024). The peer effect of persistence on student achievement. *Economics of Education Review*, *102*, Article 102574.