Journal of Economic Behavior and Organization 191 (2021) 738–751

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

Mathematics camps: A gift for gifted students?☆

Ainoa Aparicio Fenoll[a], Flavia Coda Moscarola[b], Sarah Zaccagni[c,*]

[a] *University of Turin, ESOMAS Department, Corso Unione Sovietica 218 Bis, Torino 10134, Italy*
[b] *Planning, Research and Evaluation Department of Fondazione Compagnia di San Paolo, CeRP-Collegio Carlo Alberto, Italy*
[c] *University of Copenhagen, Department of Economics and CEBI (Center for Economic Behavior and Inequality), Øster Farimagsgade 5, Building 35, Copenhagen DK-1353, Denmark*

ARTICLE INFO

Article history:
Received 23 December 2020
Revised 13 August 2021
Accepted 21 September 2021

JEL classification:
C93
D04
I21
I26
I28

Keywords:
Randomized control trial
Mathematics
Extra-curricular courses
Gifted students
Talented students

ABSTRACT

We evaluate the short-run impact of a mathematics camp for gifted high school students. During the camp, studets work in teams, trying to solve advanced mathematical problems with the help of manipulatives. We randomize participation in the camp and test the effects of such participation on problem-solving skills, personality traits, and career intentions. Results show that participants improve their problem-solving skills, especially in questions that require the use of logic. We also find positive effects on students' personality traits: students declare to be less neurotic and more extroverted. Gifted students with relatively lower school math scores benefit more from the program. Finally, participating in the mathematics camp makes students more willing to go to university.

1. Introduction

Programs for gifted students find their rationale in the public interest to promote individual self-fulfillment and in the positive returns in terms of economic and societal development (Renzulli, 2012). Since the late '80s, several special education programs targeting gifted students have been introduced in the US.[1] These programs involve around 7% of the overall

☆ We are grateful to Christina Gravert, Mette Ejrnæs, Giovanni Mastrobuoni, Ignacio Monzón, Juan Morales, Catia Nicodemo, Marco Piovesan, Silvana Maria Robone, Almudena Sevilla, Giuseppe Sorrenti, Miriam Wüst, Ingo Zettler, and two anonymous referees for their useful comments. This randomized controlled trial is registered at the AEA RCT Registry with identification number AEARCTR-0003781. The data that support the findings of this study and the code are available upon request. The authors have no competing interests, or other interests that might be perceived to influence the interpretation of the article. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We thank the Mathesis Association and Paola Sabbione (Fondazione Compagnia di San Paolo) for their support in the implementation of the experiment. We thank seminar participants at Collegio Carlo Alberto and at the TRIBE Lab, University of Copenhagen. The activities of CEBI are financed by the Danish National Research Foundation, Grant DNRF134. All errors are our own.

* Corresponding author.
E-mail addresses: ainoa.apariciofenoll@unito.it (A.A. Fenoll), flavia.codamoscarola@compagniadisanpaolo.it (F.C. Moscarola), Sarah.Zaccagni@econ.ku.dk (S. Zaccagni).

[1] The US National Association for Gifted Students regrets the absence of a uniform federal policy for "gifted services". See https://www.nagc.org/ for more detailed information.

https://doi.org/10.1016/j.jebo.2021.09.036
0167-2681/

student population (Card and Giuliano, 2014). Recently, in Europe, many countries have introduced laws to meet gifted students' needs for special intervention[2]. Still, research studies addressing the effectiveness of gifted programs focus on the US.[3] Most of the evaluated US programs are public school programs, implemented at the state or district level. US programs stand in contrast with gifted programs in Europe, which are often short-lived, extra-curricular, and up to the initiative of teachers or private institutions. One example of these European programs is the Mathesis mathematics camp for gifted students that takes place every year in northwestern Italy. This specific program is especially interesting because it replicates on a smaller scale the International Mathematical Olympiad, which involves more than 100 countries that every year send teams of up to six students to compete in math games.

We estimate the impact of the program using a randomized control trial, generally recognized as the "gold standard" method for program evaluation. We evaluate the effects of the camp not only on problem-solving skills but also on non-cognitive outcomes like self-assessed personality traits and career intentions. The math camp is a long-lasting initiative promoted by Mathesis, an association of mathematics teachers, and it is financed by Fondazione Compagnia di San Paolo, one of the largest philanthropic organizations in Italy. The camp takes place yearly at the end of the academic year (May-June) and lasts three days. It targets students of high ability from high schools in Italian northwestern regions. Participants are in high-school years one to four (ages 14 to 18). School mathematics teachers select the top-performing students in each class to attend the camp. During the camp, students are randomly assigned to teams of around six students and exposed to advanced mathematical problems that are unrelated to the school curricula. These problems must be solved in collaboration with their teammates and each team must submit one solution that is then evaluated by the teachers. The team with the best evaluation for each grade receives symbolic prizes at the end of the camp. Teachers facilitate mathematical reasoning by providing each team with mathematical manipulatives. Hence, the camp is characterized by peer-to-peer learning, "inquiry-based" activities, and a "hands-on" working style.

In this paper, we shed light on the short-run impacts of participation in the mathematics camp using a randomized control trial. Teachers typically choose a small number of students per class (the median is two) according to a subjective criterion, which heavily relies on their school mathematics score but also takes into account students' interest in mathematics and teacher-assessed potential. For our evaluation, we asked teachers to select one additional student per class, the first student they would have chosen if they needed to fill an extra slot in the camp. In other words, teachers select those students they would have selected in the absence of our evaluation and the first student on the "waiting list". We then randomly selected one out of the signaled students in each class and exclude him/her from the mathematics camp, while the remaining students participate in the camp. The group of participants in the mathematics camp constitutes our treatment group ($N = 988$), while excluded students are our control group ($N = 400$). We then compare the answers of treated and control students to a questionnaire including demographic, psychological, and career intentions questions, as well as mathematical problems. Additionally, we explore the heterogeneity of treatment effects by socio-demographic characteristics and by performance in mathematics tests at school.

Which effects do we expect to find? Evidence of the effectiveness of government interventions in education primarily focuses on US programs as the Tennessee STAR class size experiments (Krueger, 1999; Krueger and Whitmore, 2001) and the New York City school choice program (Krueger and Zhu, 2003). Outside the US, there are fewer contributions to the literature on the topic, including the Angrist and Lavy (2009) study of the Israeli high school entry exam and Machin et al. (2004) study on the impact of the Excellence in Cities (EiC) U.K. program. Focusing on evaluations of the effects of gifted student programs, previous literature finds mixed results. Murphy (2009) and Bui et al. (2014) elicit the effect of gifted or talented services supplied to US students using a regression discontinuity approach comparing admitted students just above the minimum threshold for admittance with students not admitted but just below the threshold. Both studies find little or no impact on marginal students. Dougherty et al. (2014) use the same evaluation design in the context of the Wake County Public School System and find that accelerated math track for high performers has no significant effects on standardized test scores but lowers girls' scores in middle school. Targeted math acceleration has the potential to increase college readiness among disadvantaged populations only. In contrast, Bhatt (2011) finds significant improvements in mathematics scores only, but her instrument does not pass the weak instrument test. In a meta-analysis, Cooper et al. (2000) show that summer programs for accelerated learning have positive impacts on the knowledge and skills of participants, with middle-class students benefiting more from those programs. The results are also ambiguous for the impact of mathematics courses on academic performance which is positive in some contexts (Cortes et al., 2015; Aughinbaugh, 2012) but negative in others (Clotfelter et al., 2015).

The difference in the effects of mathematics programs can be explained by differences in program characteristics. Previous literature indicates that effective mathematics programs are characterized by "inquiry-oriented" instruction (Blazar, 2015), frequent teacher feedback, the use of data to guide instruction, "high-dosage" tutoring, increased instructional time, and high expectations (Dobbie and Fryer, 2013). Ellison and Swanson (2016) finds that the quality of school and teachers are particularly important for gifted students. Fuchs et al. (1997) show that peer-to-peer learning increases mathematics achievement as measured by the Stanford Achievement Test (SAT-10). Hence, given the characteristics of our program, we expect that it positively affects mathematical problem-solving skills.

[2] For a comprehensive description of gifted education in Europe see Tourón and Freeman (2018)

[3] See Bhatt (2011); Bui et al. (2014); Dougherty et al. (2014); Card and Giuliano (2014), and Murphy (2009).

In line with our expectations, we find that students participating in the mathematics camp improve their problem-solving skills. The improvement is higher in problems that require the use of logic rather than problems that require formal mathematics knowledge (formulas, standard solving methods, etc.). Consistent with findings in Cortes et al. (2015), the estimated positive effect is heterogeneous, and it is stronger for 14-year-old students and those with high-educated parents. Card and Giuliano (2014) finds that gifted programs implemented in the largest US school districts do not affect high IQ students but have positive and relatively large effects on students with good school performance who would not normally qualify for gifted programs. In line with their finding, we also find that gifted students with relatively lower school mathematics scores benefit more from the mathematics camp.

Our paper contributes to a broader literature on the effects of extra-curricular activities on students' competencies, including both gifted and non-gifted students. This literature shows that there is a substantial degree of heterogeneity in the effects of extra-curricular activities. The initial level of competencies, family income, cultural background, and the type of activities implemented are mediating factors in the improvement in students' outcomes. Hvidman et al. (2020) analyze the effect of summer courses on students assessed "not-ready for further education after compulsory school" and find small positive effects on math competencies. Kim and Quinn (2013) and Lauer et al. (2006) find that guided reading activities conducted during summer improve reading skills for children, especially from low-income households. Lauer et al. (2006) and Mariano and Martorell (2013) and Matsudaira (2008) also find beneficial effects for several extra-curricular activities about mathematics and reading, especially if programs include tutoring. In contrast, Jacob and Lefgren (2004) study the case of Chicago High schools and find that the effect of remedial programs is almost limited to third-year attendees.

Given previous findings that personality traits affect learning strategies and performance (Blickle, 1996) and that it is important to consider personality traits to optimize the efficacy of training and cognitive interventions (Studer-Luethi et al., 2012), we are also interested in the impact of the mathematics camp on personality traits. Qualitative assessments of the impact of gifted programs highlight that students think that participating in a gifted program helps accomplish more difficult tasks which in turn increases self-esteem (Hertzog, 2003). At the same time, the effects of the camp on self-concept may be ambiguous as talented students surrounded by other talented students may update their beliefs about their position in the ability distribution upwards or downwards and affect the perception of what factors play a bigger role in determining their success (i.e. talent versus diligence). For this reason, we test empirically the impact of participation in the camp on answers to the Big Five questionnaire, which captures five aspects of personality (i.e., Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) and a series of questions about students' opinion on what determines their academic success. For the sake of time and to avoid survey fatigue (Porter et al., 2004), we use a short version of the Big Five inventory, asking the subjects to self-declare their personality traits. Recent literature shows that short versions still capture personality traits (Rammstedt and John, 2007; Soto and John, 2017; Topolewska et al., 2014). We find that the camp reduces the incidence of neuroticism and fosters the perception of being extroverted.[4]

Our finding that the camp fosters mathematical knowledge is important given the established results on the impact of mathematics on future outcomes. Aughinbaugh (2012) shows that a more rigorous high school mathematics curriculum is associated with a higher probability of attending college. Their household fixed effects results imply that students who take an advanced academic mathematics curriculum in high school (algebra II or precalculus, trigonometry, or calculus) are about 17% more likely to go to college compared to those students whose highest mathematics class was algebra I or geometry. Moreover, mathematics courses increase students' propensity to sign up for a mathematics degree (Maestri, 2013). Our results show that students attending the camp are more likely to declare that they will enroll in university but we find no effect on intentions to study a STEM degree. Studying mathematics also generates high labor market returns, especially in the early stages of working life, as a result of skill-biased technological changes (Deming and Noray, 2018).

The outline of the rest of the paper is as follows. Section 2 provides details on how the mathematics camp operates and on the design of our randomized control trial. We describe our data, the randomization, and characterize our sample in Section 3. Section 4 specifies the empirical strategy to estimate the effects of the camp. Section 5 establishes the main results on the impact of the mathematics camp. It also includes robustness checks and tests for heterogeneous treatment effects. Finally Section 6 concludes.

2. Program background and evaluation design

2.1. Description of the program

The evaluated gifted mathematics program takes place in Italy. In general, Italian students perform poorly in mathematics according to PISA. The share of Italian top performers (students who attained levels 5 or 6 in PISA tests) is only 9.5% compared to 11.4% in the entire OECD. In this context, gifted programs may be useful to improve performance at the top of the Italian distribution.

[4] Both these personality traits have been linked to performance, motivation, and psychological vulnerability in the existing literature. Studer-Luethi et al. (2012) report a significant interaction of neuroticism and training as determinants of efficacy. Lower neuroticism in adolescents is also associated with a lower risk of negative automatic thoughts and impulsivity (Kercher et al., 2009; Fetterman et al., 2010). Turban et al. (2017) explore the role of extroversion later in life and indicate a relationship of proactive personality and extroversion with career success, mediated by mentoring and organizational knowledge.

Before 2007, Italian education authorities had paid limited attention to gifted students. Mönks et al. (2005) elaborated a comprehensive review of education measures for gifted students across EU countries. In 2005, Italian talented children were allowed to start in school earlier (ISCED level 1), complete two grades in one year (ISCED levels 2 & 3), and take the final examinations in advance (ISCED level 3). Law 1/2007 and Decree 262/2007 first introduced the notion of fostering excellence in school in Italian legislation. In practice, they introduced a scholarship dedicated to top-achieving students. Only recently, in 2019, did the Ministry of Education acknowledge gifted students' needs for special attention (l.562/2019). However, Italian authorities have not promoted concrete actions until now.

In a context of low mathematics achievement and little institutional attention to gifted students, we evaluate the effectiveness of a mathematics program targeting gifted students. The Mathesis Mathematics Camp is an intensive three-day extra-curricular mathematics program that targets students from 48 high schools in the northwestern regions of Italy. Attendance is restricted to high-performing students by invitation from their high school mathematics teachers. Selection criteria are discretionary but most teachers mainly rely on previous mathematics achievements. Other criteria include students' motivation and teacher-assessed potential.

The camp is organized by Associazione Subalpina Mathesis, a well-established association of high school mathematics teachers. It has taken place yearly since 1995. The initiative is almost entirely financed by Fondazione Compagnia di San Paolo, one of the largest philanthropic organizations in Italy. Students' participation fee is about 90 euros and covers all expenses (including hotel accommodation, meals, and transportation to and from the camp). Schools pay the participation fee of low-income students.

Fondazione Compagnia di San Paolo is planning to propose to the Italian Ministry of Education the adoption of the Mathesis mathematics camp approach at the national level. They intend to extend the camp geographically while maintaining the same design (selection criteria, working method, etc.). The results of the current study are informative about the convenience of such an extension and whether targeting specific groups of students would make the camp more efficient.

The camp takes place every year between May and early June in Bardonecchia, a mountain site in the surroundings of Turin. About 1500 students from the first four years of high school (grades nine to twelve) participate in each edition. Students in the last year of high school (thirteenth grade) are excluded because they need to prepare for the high school diploma final examination. The teaching staff is composed of approximately 120 high school professors, six university professors, eight graduate students, and 20 undergraduate students. Due to location capacity constraints, students are divided into four waves in which students of each grade are equally represented. All waves are identical and last three days each.

During the mathematics camp, students work in open spaces where each space is allocated to one grade. Students are assigned to teams of six participants (exceptions of five and seven participants are allowed when necessary) who work at separate tables. Teachers circulate between the tables to solve doubts about the wording of exercises and supervise activities.

In each edition, participants in the same grade work on a given mathematic topic by solving a series of related problems with the help of manipulatives provided by teachers. For instance, in 2019, first-year students worked on algorithms, second-year students devoted their time to the concept of infinity, third-year students studied bar codes and cryptocurrencies, while four-year students focused on non-Euclidean geometry. Teachers evaluate the proposed solution not only based on its correctness but also the originality of the problem-solving method. The team with the highest accumulated score is awarded a symbolic prize.

2.2. Randomized experiment design

Our study was conducted between November 2018 and June 2019.[5] To carry out the evaluation, we requested Mathesis to implement some changes to the ordinary organization of the mathematics camps. By January 2019, the high school teachers involved in the program provided the list of participants. In a regular edition of the camp, each teacher would have chosen N_i students in each class i to participate in the camp. For the evaluation, teachers were asked to select $N_i + 1$ students. The extra student must be the one teachers would have chosen if they needed to fill an extra slot in the camp. In other words, teachers need to add the first student in the "waiting list". We would then randomly select one student per class to be excluded from participation in the camp. The set of excluded individuals forms our control group. We could have asked teachers to select only those students they would have selected in a regular year and exclude some of them but: (i) this would have significantly changed the nature of the camp experience as there would be significantly fewer students participating, (ii) Mathesis teachers opposed this alternative as it would imply wasting resources. Alternatively, we could have extended the set of students signaled by teachers (for instance, we could have asked two additional students per class to have treatment and control groups of similar size) but, as explained to us by Mathesis teachers, this alternative would make our sample too different from the population of interest, i.e., gifted students. Following our rule, teachers selected 2125 students to potentially participate in the camp.

In February 2019, teachers administered the pre-camp questionnaires to all the students in the list of potential participants at the regular class time under their supervision. The test consists of 74 questions, divided into four sections: six student identification questions, 14 socio-demographic questions, 45 psychological and aptitude questions, nine mathematics-

[5] See Appendix Appendix A for a complete timeline of the experiment.

related questions including three problems, and questions about the methods used to solve the three problems. We reproduce the pre-camp questionnaire in Appendix I in the Supplementary Materials.

By March 2019, based on the lists of 2125 candidates provided by the teachers, we randomized participation to the camp. As we used stratified randomization by class, each class is represented in the final sample by the number of students that would have participated in the camp in the absence of our evaluation. To foster collaboration in completing the post-camp questionnaires, we gave excluded students in the first, second, and third years of high school the opportunity to participate in the camp in the following year. To fourth-year students, we offered the opportunity to participate in a summer school at the University of Turin about mathematical applications to economics. After randomization, 1455 students participated in the camp, and 670 became part of the control group.

A week after the mathematics camp, teachers administered a post-camp questionnaire to students in treatment and control groups at the regular class time under their supervision. The post-camp questionnaire consists of the six student identification questions, the 14 socio-demographic questions in the pre-camp questionnaire (to be filled only by students who did not answer the pre-camp questionnaire), 50 psychological and aptitude questions (the 45 questions in the pre-camp questionnaire and the Big Five), five mathematics problems, and three questions about the methods used to solve the problems. We reproduce the post-camp questionnaire in Appendix II in the Supplementary Materials.

We measure problem-solving skills using five mathematics problems. The low number of problems allow us to avoid problems related to fatigue or disengagement (Balart et al., 2018; Balart and Oosterveen, 2019). Moreover, Bertoni et al. (2021) show that longer tests do not necessarily provide a more accurate measure of students' competencies.

3. Data and descriptive statistics

3.1. Pre-camp questionnaire and randomization

In the analysis, we use information from two questionnaires: the pre-camp questionnaire administered in February 2019 and the post-camp questionnaire administered one week after the camp in June 2019. We use students' answers to the pre-camp questionnaire to check that the groups of treated and control students are comparable ex-ante in terms of problem-solving skills and to collect information on socio-demographic characteristics that we use as controls in our regressions. We use the answers to the post-camp questionnaire to measure differences emerging between the two groups as a consequence of the camp. In this section, we present descriptive statistics for the main individual characteristics, for the whole sample as well as treatment and the control groups, separately.

Fig. 1 illustrates the number of students participating in each phase of the experiment. A total of 48 high schools and 2125 students were initially included in the experiment. A few days before the camp, 74 students declared that they were unable to attend and were replaced by students in the control group. To avoid these replacements affecting our results, we excluded all students belonging to classes in which there was a replacement. Following this approach, we excluded 255 students. Excluding those students leaves the proportion of treated students in the sample unchanged.[6] Moreover, we only included, in the analysis, students who answered both the pre-camp and the post-camp questionnaires, i.e., those for whom we have information on the outcomes of interest and the socio-demographic controls. Out of the 1870 students not affected by replacements, 1602 students answered the post-camp questionnaire and 1388 answered both questionnaires. Therefore, the final sample is composed of 1388 students: 988 in the treatment group and 400 in the control group.[7]

We assign potential camp participants to treated and control groups using stratified randomization (Athey and Imbens, 2017). Each potential participant belongs to one stratum or class i. We denote the number of classes by I. Let N_i be the number of participants in the camp who belong to class i. As teachers select one additional student per class, the number of units in stratum i, i.e., the number of potential participants in class i is $N_i + 1$. We then randomly select one student in each stratum and exclude him/her from participation in the camp. As a result, the number of students in the treatment group is $N_T = \Sigma_{i=1}^{I} N_i$ and the number of excluded students equals the number of strata $N_C = \Sigma_{i=1}^{I} 1 = I$. Table 9 in Appendix B shows the distribution of the number of signaled students per class. 226 out of 621 teachers signaled three students, 164 signaled two students, 135 signaled four students, and 96 signaled more than four.

Table 1 reports the socio-demographic characteristics of all students involved in the randomized control trial, for the pooled sample and separately for treated and control. There are slightly more males (53%) than females. The average student has only one sibling. There are slightly more students in the first and second years than in the third and fourth years of high school. This happens because teachers associated with Mathesis are more represented in earlier years. As teachers select the best math students to participate in the camp, their average math score in the first quarter is high (8.3 out of 10). Their Italian score is also relatively high but lower than the math score (7.7 out of 10) and their average score for all subjects is around 8. Regarding parental education, 8% have mothers with only compulsory education, around 44% have mothers with a high school diploma, and around 47% have mothers who have graduated from university. The rest are mothers with less than compulsory schooling. Fathers are less educated than mothers on average: 14% of fathers have only compulsory

[6] A regression of a dummy for excluded student on the treatment indicator gives a coefficient of -0.006 with p-value equal to 0.709.

[7] As a result of the exclusion of non-respondent students, 70 classes are represented by only one student in the sample. Results remain invariant when we exclude those classes from our estimations.

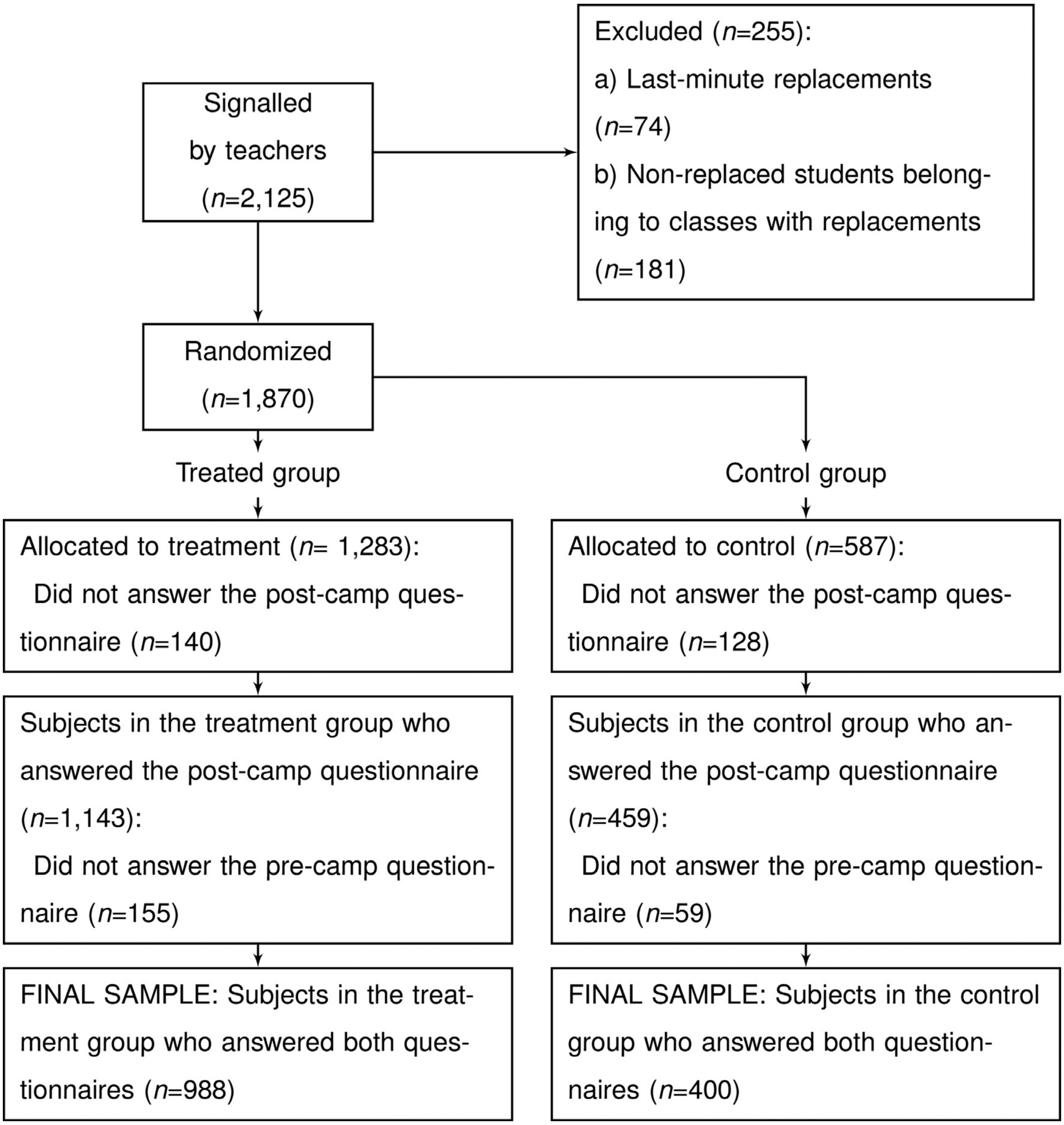


Fig. 1. Flowchart of sample definition.

schooling, almost 42% of fathers are high school graduates, while 43% have attained university degrees. Again, the rest are fathers with less than compulsory schooling. We also test whether the randomization produced homogenous groups in terms of pre-determined characteristics. In particular, we test whether treated and control students in the final sample have comparable socio-demographic characteristics and performance in the mathematical problems solved before the camp. To this end, the second column contains average values for treated students, the third column displays average values for controls, and the fourth column shows the difference between the two groups and whether this difference is significant at conventional levels. p-values show that there are no significant pre-camp differences between the two groups, except for one mathematics question ("correct parabola", p=0.08). A joint test of significance across all pre-camp variables results in a p-value of 0.6546.

We also use the answers to the pre-camp questionnaire to provide evidence that the exclusion of students, because they did not fill the post-camp questionnaire, does not affect our estimates. We show that students in the sample and those excluded because they did not fill the post-camp questionnaire are not different in terms of observable characteristics (see Table A.1 in Appendix III in Supplementary Materials). All differences are statistically insignificant at conventional levels (except for a 10% significance in the difference in the approach adopted to solve one of the three exercises).

Table 1
Descriptive statistics and balance test.

Variable	Mean pooled	Mean treated	Mean control	Difference
Male	0.527	0.560	0.540	0.020
Year of Birth	2002.64	2002.6	2002.7	−0.089
N siblings	1.074	1.067	1.093	−0.026
Class == I	0.29	0.282	0.31	−0.028
Class == II	0.251	0.248	0.26	−0.012
Class == III	0.233	0.240	0.213	0.028
Class == IV	0.226	0.229	0.218	0.012
Math score	8.302	8.203	8.267	−0.064
Italian score	7.732	7.562	7.492	0.070
Average score	8.013	7.950	7.930	0.020
Mother below high school	0.081	0.08	0.085	−0.005
Mother high school	0.444	0.447	0.433	0.015
Mother university	0.467	0.466	0.473	−0.007
Father below high school	0.14	0.147	0.123	0.025
Father high school	0.415	0.418	0.405	0.013
Father university	0.433	0.424	0.458	−0.034
Pre-test				
Pre-camp test score	2.781	2.795	2.748	0.048
Standardized pre-camp test score	0.362	0.374	0.333	0.041
Correct system of equations	0.960	0.960	0.960	0.0004
Answered through logic	0.244	0.246	0.241	0.005
Answered through system of eq.	0.648	0.644	0.658	−0.014
Answered through attempts	0.039	0.036	0.046	−0.010
Correct parabola	0.851	0.862	0.825	0.037*
Answered through formula	0.821	0.815	0.836	−0.021
Answered through attempts	0.051	0.049	0.055	−0.005
Correct rectangle	0.970	0.973	0.963	0.010
Answered through formula	0.326	0.332	0.313	0.019
Guesses drawing	0.124	0.131	0.108	0.023
Guesses without drawing	0.467	0.459	0.485	0.026

Note: Data is from the pre-camp questionnaire complemented with data from the post-camp questionnaire when missing. The number of observations is 1388, 988 in the treatment group and 400 in the control group. The differences are marked with * if the level of significance is between 5% and 10%.

Table 2
Descriptive statistics: post-camp problem solving skills .

Variable	Mean	Std. Dev.	Min.	Max.
Math score	4.563	0.749	1	5
Correct system of equations	0.968	0.175	0	1
Correct parabola	0.888	0.315	0	1
Correct trapezoid	0.955	0.208	0	1
Correct Logic I	0.854	0.353	0	1
Correct Logic II	0.897	0.304	0	1
N	1388			
Treatment group:				
Math score	4.62	0.691	1	5
Correct system of equations	0.976	0.154	0	1
Correct parabola	0.902	0.298	0	1
Correct trapezoid	0.962	0.192	0	1
Correct Logic I	0.871	0.335	0	1
Correct Logic II	0.91	0.286	0	1
N	988			
Control group:				
Math score	4.42	0.86	1	5
Correct system of equations	0.950	0.218	0	1
Correct parabola	0.855	0.353	0	1
Correct trapezoid	0.938	0.242	0	1
Correct Logic I	0.813	0.391	0	1
Correct Logic II	0.865	0.342	0	1
N	400			

Note: Data is from the post-camp questionnaire.

Table 3
Descriptive statistics: post-camp big five traits.

Variable	Mean	Std. Dev.	Min.	Max.
Complete sample:				
Do you consider yourself friendly?	7.983	1.532	1	10
Do you consider yourself neurotic?	6.189	2.557	1	10
Do you consider yourself conscientious and responsible?	8.300	1.45	1	10
Do you consider yourself extroverted and sociable?	7.262	2.013	1	10
Do you consider yourself open-minded?	8.579	1.462	1	10
N	1388			
Treatment group:				
Do you consider yourself friendly?	8.027	1.503	1	10
Do you consider yourself neurotic?	6.084	2.527	1	10
Do you consider yourself conscientious and responsible?	8.306	1.444	1	10
Do you consider yourself extroverted and sociable?	7.354	1.963	1	10
Do you consider yourself open-minded?	8.589	1.44	1	10
N	988			
Control group:				
Do you consider yourself friendly?	7.875	1.599	1	10
Do you consider yourself neurotic?	6.45	2.615	1	10
Do you consider yourself conscientious and responsible?	8.288	1.468	1	10
Do you consider yourself extroverted and sociable?	7.033	2.115	1	10
Do you consider yourself open-minded?	8.553	1.516	1	10
N	400			

Note: Data is from the post-camp questionnaire.

Table 2 summarizes the answers to the five mathematics problems administered in the post-camp questionnaire. We show descriptive statistics for the complete sample, treatment, and control groups, separately. The first three questions are standard mathematical problems: the first is a problem that can be solved through a system of equations, the second regards a second-order equation, and the third is about geometry (in particular, a trapezoid). The last two questions do not require the use of mathematical concepts, they can be solved using logic. We include the latter to test how students react when they feel out of their "mathematics comfort zone". We call the sum of correct answers to the five questions *Math score*.

We designed the mathematical problems in a way that replicates the school score obtained by potential participants in the camp.[8] The average Math score is 4.5 out of 5 which is equivalent to 9 out of 10 in the school score. This figure is slightly higher than the average school scores obtained in the first quarter (8.3). However, Mathesis teachers argue that first-quarter school scores are lower because many teachers are stricter in intermediate evaluations to motivate students. Teachers who adopt this practice give the deserved (often higher) final scores at the end of the academic year. As reported by Mathesis, 9 is a reasonable average final score for students in our sample. Participants to the camp answer 0.2 more problems correctly than excluded students, corresponding to an increase of 4.5% in the Math score. Treated students perform better than controls in all problems. The highest differences between treated and controls appear in the problems that require the use of logic. In the next section, we quantify these differences using regressions.

In Table 3, we describe students' Big Five personality traits that we use as additional dependent variables. Again, we show descriptive statistics for the entire sample, the treatment group, and the control group, separately. The Big Five model classifies personality traits into five categories: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Rothmann and Coetzer, 2003). The most prevalent personality trait among students in our sample is open-mindedness: the average score is 8.6 out of 10. After being open-minded, students decreasingly declare to be conscientious and responsible, friendly, extroverted and sociable, and finally neurotic. Students in the treatment group appear slightly friendlier, less neurotic, and more extroverted than students in the control group. We find no differences in responsibility or open-mindedness between treated and control students. We test whether these differences are statistically relevant using regressions in the next section.

4. Econometric strategy

The randomized design is our source of identification. We estimate the average treatment effect of participating in the camp by estimating the following model by ordinary least squares:

$$Y_i = \beta_1 T_i + \beta_2 \mathbf{X}_i + class_i + \epsilon_i, \tag{1}$$

where Y_i is one of our outcome variables measuring student i's problem-solving skills, psychological traits, or career intentions. Regarding problem-solving skills, we study the number of correct answers (math score) and dummies for having solved each problem correctly. In the analysis of psychological traits, we consider the Big Five personality traits. Finally, the outcomes that measure career intentions comprise intentions to go to university and to study for a STEM degree. T_i is a

[8] We designed the Mathematics problems and asked Mathesis teachers to check that the problems were adequate for our study.

Table 4
Effect of treatment on problem-solving abilities.

	Math score (1)	Std. score (2)	Sys. of eq. (3)	Parabola (4)	Trapezoid (5)	Logic I (6)	Logic II (7)
Treatment	0.194	0.262	0.018	0.042	0.035	0.052	0.047
	(0.045)***	(0.063)***	(0.012)	(0.019)**	(0.015)**	(0.021)**	(0.02)**
RW p-values	.	.	0.164	0.142	0.129	0.142	0.142
Obs.	1388	1388	1388	1388	1388	1388	1388
R^2	0.031	0.03	0.009	0.016	0.015	0.027	0.018

Notes: Data is from the post-camp and the pre-camp questionnaires. The dependent variables are the math score, the standardized math score, and dummies for having answered each of the five mathematics problems correctly. All regressions include a set of control variables (gender, indicators for number of siblings, parental education dummies, and class fixed effects). The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%. Standard errors are clustered at the class level.

dummy equal to 1 if student i was randomly assigned to the treatment group (i.e., participated in the camp) and 0 otherwise. $\mathbf{X}_i$ is a vector of control variables (gender, indicators for the number of siblings, parental education dummies). *class* is a vector of class fixed effects. Finally, ϵ_i is the error term. The estimated coefficient of the treatment dummy β_1 yields the average treatment effect on the treated. All standard errors are clustered at the class level, which is the strata in our stratified randomization (de Chaisemartin and Ramirez-Cuellar, 2020).

5. Results

In this section, we discuss the results of estimating the impact of the mathematics camp on mathematical problem-solving skills, personality traits, and academic career intentions as in Eq. (1). Table 4 presents our estimates of the effect of the camp on problem-solving skills, measured by students' answers to the five mathematical problems proposed in the post-camp questionnaire. The outcome in column 1 is the raw math score which is the number of correct answers given to the five problems. In column 2, we standardize the raw math score by grade so that standardized scores in each grade have an average of zero and a standard deviation equal to one. The standardized score takes into account that different grades have different raw test scores distributions because students in different grades differ in terms of age, time spent at school, and mathematical knowledge. Columns 3–7 display the results for each of the five mathematical problems, separately. The dependent variable in column 3 equals one if the student has correctly solved the proposed system of equations; in column 4 the outcome variable equals one for students who have correctly identified a second-degree polynomial with the corresponding parabola; column 5 presents the impact of the camp on the probability of solving a geometry problem about a trapezoid; finally, columns 6 and 7 present our results for students' capacity to solve problems using logic (see Appendix II in Supplementary Materials - Post-camp questionnaire - questions 7–15).

There is a significant positive effect of the mathematics camp on mathematical problem-solving skills as measured both by the raw math score (0.194) and by the standardized math score (0.262). The estimated treatment effect ranges between 0.018 and 0.042 for the probability of solving the first three problems, which can be solved using standard mathematics tools. The camp increases the probability of solving the two problems that require logic by 0.052 and 0.047, respectively. Hence, the camp improves performance in both traditional mathematics problems and in problems that benefit from brightness and perseverance but do not require specific mathematics knowledge (formulas, rules, etc.). However, the camp is much more effective for the latter. Results remain invariant when we control for students' answers to the pre-camp mathematics problems. In addition to the standard clustered standard errors, we computed the Romano and Wolf (2005) p-values, which account for multiple hypothesis testing. They show that our estimates for the single mathematical problems are close-to but not significant at conventional levels (p-values range from 0.129 to 0.164). We also computed the heteroskedasticity robust (non clustered) standard errors and, as Table A.2 in Appendix IV in Supplementary Materials shows, they are even lower.

Our point estimates are not affected by the inclusion of controls because, as a result of randomization, socio-demographic characteristics are balanced for treated and control individuals.[9] Table A.4 in Appendix VI in Supplementary Materials shows the results for the specification excluding all controls which allows us to include also students who did not answer the pre-treatment test and who did not provide this additional information in the post-treatment test. Results remain arguably invariant. The same happens when we include students who did not answer the pre-treatment test by including indicators for missing values in the control variables. Results are displayed in Table A.5 in Appendix VI in Supplementary Materials. This is consistent with the fact that the proportion of treated students in the sample used to construct these tables is equivalent to the one in the main specification.[10]

[9] Refer to Table A.3 in Appendix V in Supplementary Materials for the coefficients associated with controls.

[10] A regression of a dummy for student who did not respond to the pre-treatment questionnaire on the treatment indicator gives a coefficient of −0.004 with a p-value of 0.852.

Table 5
Heterogeneity in the effect of treatment on math score.

	Math score				
	(1)	(2)	(3)	(4)	(5)
Treatment		0.18 (0.066)***	0.957 (0.365)***		
Treatment in year I	0.296 (0.095)***				
Treatment in year II	0.161 (0.082)**				
Treatment in year III	0.116 (0.083)				
Treatment in year IV	0.172 (0.092)*				
Treatment by male		0.028 (0.095)			
Treatment by math score			−0.092 (0.043)**		
Treatment by mother $<$ HS				0.16 (0.15)	
Treatment by mother $=$ HS				0.15 (0.064)**	
Treatment by mother $>$ HS				0.254 (0.073)***	
Treatment by father $<$ HS					−0.057 (0.109)
Treatment by father $=$ HS					0.248 (0.073)***
Treatment by father $>$ HS					0.241 (0.068)***
Obs.	1388	1388	1366	1388	1388
R^2	0.034	0.031	0.044	0.034	0.039

Notes: Data is from the post-camp and the pre-camp questionnaires. The dependent variable is the math score. All regressions include a set of control variables (gender, indicators for number of siblings, parental education dummies, and class fixed effects). The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%. Standard errors are clustered at the class level.

As explained above, the number of potential camp participants varies across classrooms. Regardless of the number of potential participants, all students but one were randomly selected to take part in the camp. As a result, the probability of treatment assignment varies across classrooms. Students in classes with two potential participants had a probability of treatment equal to 50% whereas students in classes with more participants have a probability of treatment equal to 75% or above. In additional regressions, we weight observations by these treatment probabilities. Estimates of the effect of the camp, shown in Table A.6 in Appendix VII in Supplementary Materials, become slightly stronger.

In Table 5, we explore whether there are differences in the effect of the camp on problem-solving skills across students. In particular, we analyze whether the effectiveness of the camp differs by grade (from grade nine to twelve), by gender, by math school score in the first quarter, and by parental education. In column 1, we interact the treated dummy with dummies for each of the four high-school years; in column 2, we include the interaction of the treatment dummy and a male dummy; we multiply the treated dummy with the math school score in column 3; finally, in columns 4 and 5, we explore the heterogeneity of the treatment effect by maternal and paternal education, respectively.

In column 1, we find that treatment effects are statistically different from zero for all high-school years except the third. The effect is highest for students in first year of high school (0.296) compared to the second and fourth years (0.161 and 0.172, respectively). We cannot detect differences in camp effects by gender in column 2. The negative and significant coefficient of the interaction of the treated dummy and math school score in column 3 indicates that students with lower math scores benefit more from the camp. One extra point in the math school score reduces the impact of the camp by −0.092. Columns 4 and 5 show that the impact of the camp rises with parental education, both for mothers and fathers. Unfortunately, we do not have statistical power to detect whether the coefficients in column 1 of Table 5 are significantly different from each other (all tests of equality of coefficients have p-values above 0.18).

Table 6 presents our estimates for the impact of the camp on the Big Five personality traits. We only find significant effects of the camp on neuroticism and extroversion. The camp reduces the declared level of neuroticism by −0.456 points on a 1 to 10 scale and increases self-assessed extroversion by 0.325 points on a 1 to 10 scale. The coefficients associated with the rest of the personality traits (agreeableness, conscientiousness, and openness) are positive but imprecisely estimated. Adjusting our estimated standard errors for multiple hypothesis testing confirms that the camp affects both neuroticism and extroversion. We conclude that the camp has a positive impact on student self-perception in the short run. Unfortunately,

Table 6
Effect of treatment on self-perception.

	Agreeableness (1)	Neuroticism (2)	Conscientiousness (3)	Extroversion (4)	Openess (5)
Treatment	0.124	−0.456	−0.022	0.325	0.068
	(0.104)	(0.152)***	(0.095)	(0.143)**	(0.091)
RW *p*-values	0.235	0.076	0.869	0.044	0.869
Obs.	1388	1388	1388	1388	1388
R^2	0.026	0.079	0.038	0.023	0.033

Notes: Data is from the post-camp and the pre-camp questionnaires. The dependent variables are the Big Five personality traits. All regressions include a set of control variables (gender, indicators for number of siblings, parental education dummies, and class fixed effects). The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%. Standard errors are clustered at the class level.

Table 7
Effect of treatment on determinants of school performance and self-esteem.

	Effort (1)	Talent (2)	Luck (3)	Relative intelligence (4)	Relative math skills (5)
Treatment	0.106	0.209	0.089	−0.968	−0.124
	(0.144)	(0.138)	(0.159)	(1.345)	(1.311)
Average outcome	5.83	5.49	2.88	33.18	28.30
RW *p*-values	0.717	0.550	0.717	0.442	0.781
Obs.	1379	1379	1369	1327	1327
R^2	0.042	0.029	0.014	0.074	0.072

Notes: Data is from the post-camp and the pre-camp questionnaires. The dependent variables capture to what extent students believe that their academic achievements are explained by effort, talent, and luck, respectively in columns (1)–(3). In columns (4) and (5) outcomes capture individual self-esteem measured by questions 32 and 33 in the post-camp questionnaire. All regressions include a set of control variables (gender, indicators for number of siblings, parental education dummies, and class fixed effects). The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%. Standard errors are clustered at the class level.

we cannot assess whether self-perception translates into actual changes in personality or whether the estimated effects persist over time.

We explore the heterogeneity of the effects of the camp on personality traits in Tables A.7 and A.8 in Appendix VIII in Supplementary Materials. Both effects are stronger for younger students and students with highly educated parents. While the negative impact of the camp on self-assessed neuroticism is led by females, the impact of the camp on self-assessed extroversion is led by males.

Regarding the psychological sphere, we also asked students to declare the importance of three factors, namely effort, talent, and luck, in determining their academic achievement. As shown in Table 7, we could not find effects of participation in the mathematics camp on students' likelihood to declare that their academic performance is the result of talent, effort, or luck. We also analyzed whether the camp affects self-esteem as measured by students' answers to the following two questions: (i)"If you are around 100 people who are the same age as you, how many do you usually consider more intelligent than you" and (ii) "If you are around 100 people who are the same age as you, how many do you usually consider better than you in math?". Although the sign of the coefficients indicates that the camp improves self-esteem, coefficients are not significant.

Finally, we analyze the impact of the camp on academic career intentions in Table 8. Column 1 shows that attending the camp increases the probability that students declare their intention to enroll in university by 0.7. This positive effect is led by first-year students, who increase their declared intentions to enroll in university by 0.11. Estimates remain significant after accounting for multiple hypothesis testing. The absence of camp effects on higher grade students' intentions to go to university can be explained if intentions become less malleable as the actual decision to enroll in university approaches. We do not find any gender difference in the impact of the camp on career intentions. The effect of the camp on intentions to enroll in STEM degrees is insignificant.

Previous literature often uses local estimation strategies like regression discontinuity design or instrumental variables. Instead, our paper uses a randomized control trial, which allows us to improve on existing research in terms of external validity. External validity may be damaged by the fact that adding an extra student per class for evaluation purposes may

Table 8
Effect of treatment on academic intentions.

	University (1)	STEM (2)	University (3)	STEM (4)	University (5)	STEM (6)
Treatment	0.067	0.02			0.039	−0.010
	(0.029)**	(0.029)			(0.043)	(0.043)
Treatment in year I			0.113	−0.019		
			(0.055)**	(0.047)		
Treatment in year II			0.022	−0.009		
			(0.06)	(0.057)		
Treatment in year III			0.081	0.076		
			(0.067)	(0.074)		
Treatment in year IV			0.045	0.047		
			(0.048)	(0.06)		
Treatment by male					0.054	0.057
					(0.062)	(0.067)
Average outcome	0.440	0.487	0.440	0.487	0.440	0.487
RW *p*-values	0.064	0.594	.	.	0.219	0.893
RW *p*-values (year I)	.	.	0.026	0.732	.	.
RW *p*-values (year II)	.	.	0.655	0.866	.	.
RW *p*-values (year III)	.	.	0.115	0.232	.	.
RW *p*-values (year IV)	.	.	0.400	0.474	.	.
RW *p*-values (by male)	.	.	.	.	0.604	0.593
Obs.	1384	1382	1384	1382	1384	1382
R^2	0.037	0.017	0.039	0.019	0.038	0.018

Notes: Data is from the post-camp and the pre-camp questionnaires. The dependent variables are intentions to go to university and to enroll on a STEM university degree. All regressions include a set of control variables (gender, indicators for number of siblings, parental education dummies, and class fixed effects). The coefficients are marked with * if the level of significance is between 5% and 10%, ** if the level of significance is between 1% and 5% and *** if the level of significance is less than 1%. Standard errors are clustered at the class level.

decrease the average ability of students. However, as declared by Mathesis teachers and as observed by us, the added students are highly comparable to the originally included students.[11]

6. Conclusions

Targeted programs for gifted students are important because top-performing students are likely to be tomorrow's leaders. Jeff Bezos, founder and chief executive officer of Amazon.com and richest man in the world, graduated valedictorian of his high school and summa cum laude from Princeton University. Bill Gates, founder of Microsoft and second richest man in the world, was a National Merit Scholar and scored 1590 out of 1600 on the SAT (Forbes, 2019). Also according to Forbes (2000), most top CEOs excelled in education: 24% of top CEOs in Europe have Ph.D. degrees, while the proportion of Ph.D. graduates among top CEOs in China is 33%. Many parallels exist between the characteristics used to define an effective leader and the characteristics used to describe a gifted individual. Effective leaders and gifted students are highly verbal, socially sensitive, visionary, problem solvers, critical thinkers, creative, initiators, responsible, and flexible. Although the need for more effective leaders is clear, and gifted students typically possess the characteristics to become effective leaders, gifted education in youth is often neglected and little research is devoted to it (Matthews, 2004).

In this paper, we use a randomized control trial to evaluate the impact of a gifted mathematics program on problem-solving skills, self-perception, and career intentions. The camp is representative of gifted programs in Europe because it constitutes a local, short, extra-curricular, and privately driven initiative. It also replicates the International Mathematical Olympiad that involves teams of six students from more than 100 countries, representing over 90% of the world's population, in each of its yearly editions since 1980. Its design is characterized by peer-to-peer learning, "inquiry-oriented" activities, and a "hands-on" learning style: students work in teams of approximately six students, trying to solve mathematical problems with the help of manipulatives.

In our randomized control trial, we asked teachers to select one additional student per class. This student must be the one teachers would have chosen if they needed to fill an extra slot in the camp, i.e., the first student on the "waiting list". We then randomly excluded one of the listed students in each class from participating in the mathematics camp. We then estimated the impact of the mathematics camp by comparing the answers of treated and control students on a questionnaire administered one week after the camp.

Our findings show that the mathematics camp fosters students' problem-solving skills, especially for those problems that require the use of logic rather than mastering mathematics formulas or rules. Camp participants are in years one to four of

[11] Ideally, we would have tested whether students who would have been selected in a regular year and those "in the waiting list" were comparable in terms of pre-determined characteristics. Unfortunately, teachers did not provide this information.

high school (ages 14 to 18). The treatment is particularly effective for students in the first year. We also find positive effects on students' personality traits in the short run. Students participating in the camp declare themselves to be less neurotic and more extroverted.

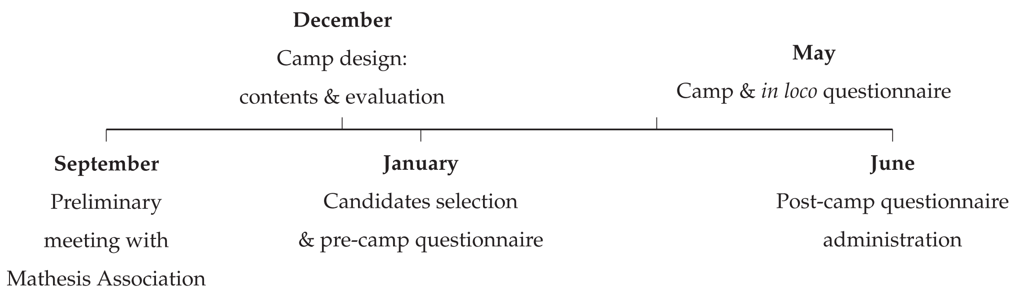
Short-run effects are relevant in our context in light of the *impressionable years hypothesis*. This hypothesis states that core attitudes, beliefs, and values crystallize during a period of great mental plasticity in early adulthood (the so-called impressionable years) and remain largely unaltered thereafter (Giuliano and Spilimbergo, 2014). These short-run effects could weaken as the camp experience becomes more distant or they may amplify if students modify their behavior in standard mathematics classes afterward. Testing this is an avenue for future research. Regarding career intentions, the camp fosters first-year students' intentions to go to university.

We confirm the findings of previous studies that effective mathematics programs are characterized by "inquiry-oriented" instruction (Blazar, 2015), frequent teacher feedback, the use of data to guide instruction, "high-dosage" tutoring, increased instructional time, and high expectations (Dobbie and Fryer, 2013).[12]

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. Timeline of the experiment



Appendix B. Number of signalled students per class

Table 9
Number of students eligible by cluster.

Signalled students per class	Number of classes	Number of students
2	164	328
3	226	678
4	135	540
5	62	310
6	11	66
7	6	42
8	8	64
9	2	18
10	4	40
11	1	11
12	0	0
13	1	13
14	0	0
15	1	15
Total	621	2125

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2021.09.036

[12] Unfortunately, in this specific setting, we do not have a way to disentangle the influence of each factor as all camp participants undergo the same experience.

References

Angrist, J., Lavy, V., 2009. The effects of high stakes high school achievement awards: evidence from a randomized trial. Am. Econ. Rev. 99 (4), 1384–1414.

Athey, S., Imbens, G.W., 2017. The econometrics of randomized experiments. In: Handbook of Economic Field Experiments, 1. Elsevier, pp. 73–140.

Aughinbaugh, A., 2012. The effects of high school math curriculum on college attendance: evidence from the NLSY97. Econ. Educ. Rev. 31 (6), 861–870.

Balart, P., Oosterveen, M., 2019. Females show more sustained performance during test-taking than males. Nat. Commun. 10 (1), 1–11.

Balart, P., Oosterveen, M., Webbink, D., 2018. Test scores, noncognitive skills and economic growth. Econ. Educ. Rev. 63, 134–153.

Bertoni, M., Brunello, G., De Benedetto, M.A., De Paola, M., 2021. Does monitoring deter future cheating? The case of external examiners in italian schools. Econ. Lett. 201, 109742.

Bhatt, R., 2011. A review of gifted and talented education in the United States. Educ. Finance Policy 6 (4), 557–582.

Blazar, D., 2015. Effective teaching in elementary mathematics: identifying classroom practices that support student achievement. Econ. Educ. Rev. 48, 16–29.

Blickle, G., 1996. Personality traits, learning stratigies, and performance. Eur. J. Personal. 10 (5), 337–352.

Bui, S.A., Craig, S.G., Imberman, S.A., 2014. Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students. Am. Econ. J. 6 (3), 30–62.

Card, D., Giuliano, L., 2014. Does Gifted Education Work? For Which Students? Technical Report. National Bureau of economic research.

de Chaisemartin, C., Ramirez-Cuellar, J., 2020. At What Level Should One Cluster Standard Errors in Paired Experiments, and in Stratified Experiments with Small Strata? Technical Report. National Bureau of Economic Research.

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2015. The aftermath of accelerating algebra evidence from district policy initiatives. J. Hum. Resour. 50 (1), 159–188.

Cooper, H., Charlton, K., Valentine, J.C., Muhlenbruck, L., Borman, G.D., 2000. Making the most of summer school: a meta-analytic and narrative review. Monogr. Soc. Res. Child Dev. i–127.

Cortes, K.E., Goodman, J.S., Nomi, T., 2015. Intensive math instruction and educational attainment long-run impacts of double-dose algebra. J. Hum. Resour. 50 (1), 108–158.

Deming, D.J., Noray, K.L., 2018. Stem Careers and the Changing Skill Requirements of Work. Technical Report. National Bureau of Economic Research.

Dobbie, W., Fryer Jr., R.G., 2013. Getting beneath the veil of effective schools: evidence from new York city. Am. Econ. J. 5 (4), 28–60.

Dougherty, S., Goodman, J., Hill, D., Schools, W. C. P., Litke, E., Page, L., 2014. Middle school math acceleration, college readiness and gender: regression discontinuity evidence from wake county, North Carolina.

Ellison, G., Swanson, A., 2016. Do schools matter for high math achievement? Evidence from the american mathematics competitions. Am. Econ. Rev. 106 (6), 1244–1277.

Fetterman, A.K., Robinson, M.D., Ode, S., Gordon, K.H., 2010. Neuroticism as a risk factor for behavioral dysregulation: a mindfulness-mediation perspective. J. Soc. Clin. Psychol. 29 (3), 301–321.

Fuchs, D., Fuchs, L.S., Mathes, P.G., Simmons, D.C., 1997. Peer-assisted learning strategies: making classrooms more responsive to diversity. Am. Educ. Res. J. 34 (1), 174–206.

Giuliano, P., Spilimbergo, A., 2014. Growing up in a recession. Rev. Econ. Stud. 81 (2), 787–817.

Hertzog, N.B., 2003. Impact of gifted programs from the students' perspectives. Gifted Child Q. 47 (2), 131–143.

Hvidman, C., Koch, A.K., Nafziger, J., Nielsen, S.A., Rosholm, M., 2020. An Intensive, School-Based Learning Camp Targeting Academic and Non-Cognitive Skills Evaluated in a Randomized Trial. Technical Report. IZA Discussion Papers.

Jacob, B.A., Lefgren, L., 2004. The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in Chicago. J. Hum. Resour. 39 (1), 50–79.

Kercher, A.J., Rapee, R.M., Schniering, C.A., 2009. Neuroticism, life events and negative thoughts in the development of depression in adolescent girls. J. Abnorm. Child Psychol. 37 (7), 903–915.

Kim, J.S., Quinn, D.M., 2013. The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: a meta-analysis of classroom and home interventions. Rev. Educ. Res. 83 (3), 386–431.

Krueger, A.B., 1999. Experimental estimates of education production functions. Q. J. Econ. 114 (2), 497–532.

Krueger, A.B., Whitmore, D.M., 2001. The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from project star. Econ. J. 111 (468), 1–28.

Krueger, A.B., Zhu, P., 2003. Principal stratification approach to broken randomized experiments: a case study of school choice vouchers in new york city [with comment]. J. Am. Stat. Assoc. 98 (462), 314–318.

Lauer, P.A., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D., Martin-Glenn, M.L., 2006. Out-of-school-time programs: a meta-analysis of effects for at-risk students. Rev. Educ. Res. 76 (2), 275–313.

Machin, S., McNally, S., Meghir, C., 2004. Improving pupil performance in english secondary schools: excellence in cities. J. Eur. Econ. Assoc. 2 (2–3), 396–405.

Maestri, V., 2013. Promoting scientific faculties: does it work? Evidence from Italy. Econ. Educ. Rev. 32, 168–180.

Mariano, L.T., Martorell, P., 2013. The academic effects of summer instruction and retention in new York city. Educ. Eval. Policy Anal. 35 (1), 96–117.

Matsudaira, J.D., 2008. Mandatory summer school and student achievement. J. Econom. 142 (2), 829–850.

Matthews, M.S., 2004. Leadership education for gifted and talented youth: areview of the literature. J. Educ. Gifted 28 (1), 77–113.

Mönks, F.J., Pflüger, R., Nijmegen, R.U., 2005. Gifted Education in 21 European Countries: Inventory and Perspective. Radboud University Nijmegen Nijmegen.

Murphy, P. R., 2009. Essays on gifted education's impact on student achievement.

Porter, S.R., Whitcomb, M.E., Weitzer, W.H., 2004. Multiple surveys of students and survey fatigue. New Dir. Inst. Res. 2004 (121), 63–73.

Rammstedt, B., John, O.P., 2007. Measuring personality in one minute or less: a10-item short version of the big five inventory in English and German. J. Res. Personal. 41 (1), 203–212.

Renzulli, J.S., 2012. Reexamining the role of gifted education and talent development for the 21st century: a four-part theoretical approach. Gifted Child Q. 56 (3), 150–159.

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

Rothmann, S., Coetzer, E.P., 2003. The big five personality dimensions and job performance. SA J. Ind. Psychol. 29 (1), 68–74.

Soto, C.J., John, O.P., 2017. Short and extra-short forms of the big five inventory–2: the BFI-2-S and BFI-2-XS. J. Res. Personal. 68, 69–81.

Studer-Luethi, B., Jaeggi, S.M., Buschkuehl, M., Perrig, W.J., 2012. Influence of neuroticism and conscientiousness on working memory training outcome. Personal. Individ. Differ. 53 (1), 44–49.

Topolewska, E., Skimina, E., Strus, W., Cieciuch, J., Rowiński, T., 2014. The short IPIP-BFM-20 questionnaire for measuring the big five. Rocz. Psychol. 17 (2), 385–402.

Tourón, J., Freeman, J., 2018. Gifted education in europe: Implications for policymakers and educators..

Turban, D.B., Moake, T.R., Wu, S.Y.-H., Cheung, Y.H., 2017. Linking extroversion and proactive personality to career success: the role of mentoring received and knowledge. J. Career Dev. 44 (1), 20–33.