TACKLING THE GENDER GAP IN MATHEMATICS WITH INNOVATIVE TEACHING METHODS

Di Tommaso M.L., Contini D., Bernardi M., De Rosa D., Ferrara F., Piazzalunga D., Robutti O.

Background

According to the last PISA assessment (OECD 2016), Italy is one of the countries with the highest gender gap in math (GGM) at age 15. Similar results are shown for children in fourth grade in the 2015 TIMMS study (Mullis et al. 2016). Contini et al. (2017)[1], using the National Assessment for Italy, show that boys outperform girls in mathematics from 2nd to 10th grade. Similar to other countries[2], the gender gap in math is increasing with age and is largest at high percentiles of the ability distribution. The evidence from most assessments is that the gender gap in math is particularly high in the area of numeracy. The presence of a substantial females' disadvantage in math is of particular importance, because it is likely to be a cause of the critically low share of women choosing STEM (Science Technology Engineering and Mathematics) disciplines at university, of gender segregation in the labor market, and gender pay gaps[3].

One of the several explanations proposed for the existence of the GGM points to the role of educational methods and practices. Some studies suggest that when mathematics' teaching is centered upon problem solving, involving students in discussions and investigative work as opposed to traditional passive methods – more common in Italy – the gender gap in math decreases and can even disappear[4]. These researchers frame the problems of the GGM within the consolidated stream of 'constructivist and social' methods (CSM) in mathematical teaching/learning[5]. In a nutshell, according to CSM, mathematical learning involves activity on the part of the learner, leading to the idea that learners 'make things' together and 'communities of practice' are created[6]. The focus is more on participation than on passive knowledge acquisition.

Purpose

The hypothesis is that CSM type teaching in mathematics will contribute to reduce the GGM. The aim of the research is to test whether CSM-type teaching methods (described above) applied to both girls and boys reduce the GGM.

---

[1] **Contini D., Di Tommaso M.L., Mendolia S. (2017),** "The gender gap in mathematics achievement: Evidence from Italian data." Economics of Education Review, 58, pp 32-42.

[2] **Fryer, Roland G., and Steven D. Levitt. 2010**. "An Empirical Analysis of the Gender Gap in Mathematics." American Economic Journal: Applied Economics, 2(2): 210-40.

[3] **European Commission 2006**, Women in Science and Technology- the Business Perspective, Luxembourg: Office for Official Publication of the European Communities . **European Commission 2012**, Enhancing excellence, gender equality and efficiency in research and innovation, Luxembourg: Office for Official Publication of the European Community. **European Commission 2015**. Science is a girls' thing, Newsletter Nov. 2015. **National Academy of Science (2007)** Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering.

[4] **Boaler, J. (2002a).** The development of disciplinary relationships: Knowledge, practice and identity in mathematics classrooms. For the learning of mathematics 22(1), p. 42-47. **Zohar, A., & Sela, D. (2003).** Her physics, his physics: gender issues in Israeli advanced placement physics classes." International Journal of Science Education, 25(2), 245-268. **Boaler, J. (2002b**). Experiencing School Mathematics: Traditional and Reform Approaches to Teaching and Their Impact on Student Learning. Mahwah, NJ, Lawrence Erlbaum Association. **Boaler, J. (2009**). The Elephant in the Classroom: Helping Children Learn and Love Maths. Souvenir Press: London. **Boaler, J. & Greeno, J. (2000).** Identity, Agency and Knowing in Mathematics Worlds. in J. Boaler (Ed) Multiple Perspectives on Mathematics Teaching and Learning. Westport, CT:Ablex Publishing (pp171-200), **OECD (2016)**, *PISA 2015 Results (Volume I): Excellence and Equity in Education*, OECD Publishing, Paris.

[5] For a survey see **Gutierrez, A., & Boero, P. (2006).** Handbook of Research on the Psychology of Mathematics Education Past, Present and Future. Rotterdam: Sense publ. pp. 305-428).

[6] **Lave, J. & Wenger, E. (1991).** Situated Learning. Legitimate peripheral participation. Cambridge: University of Cambridge Press.

Setting The research took place in 25 schools in Italy, in the Torino province, in 2019. It involved approximately 1,000 children in 50 third grade classes. We chose third grade classes because previous research[7] showed that the GGM starts in 2nd grade. The laboratories took place in the period between March and May 2019.

Intervention

The intervention – designed by scholars in the field of math education – consists in five laboratory sessions of three hours each, in five consecutive weeks, taking place at the classroom level. The intervention is carried out by 4 trained tutors with a math education background. School teachers remain in the classroom with the role of observers. Children are divided in small groups and are asked to do group activities and some individual work. All students take part to the activities, including student with disabilities or special needs (however, their data will not be processed). Children selected in the control group follow the usual curricula.

Research Design:

The impact of the intervention if evaluated by means of a randomized control trial (RCT). The RCT intervention involves approximately 1000 third grade pupils of 50 classes belonging to 25 schools. Schools participated voluntarily with at least two classes: one was randomized to treatment and one to the control group. To avoid contamination and confounding effects, we required the two classes to have different math teachers and not to be involved in other math extra-curricular activities in the current schooling year.

Analysis

All children in the treatment and control groups, sat a pre-test on math numeracy 2 weeks before the start of the intervention. The test – specifically designed by the math scholars of the research group with a similar conceptual framework of the national math standardized assessment, and previously tested within a pilot study with IRT methods – was administered in the classrooms under the supervision of the tutors in charge of the labs. Similarly, a post test was administered approximately one month after the end of the lab in each school. Tests were graded blindly by the tutors.

Results:

A substantial gender gap was existing before the intervention took place. On average, girls scored 0.21 sd. less than boys (variability was similar in the two groups).

Balancing tests

We conducted balancing tests with respect to children characteristics (pre-test score, presence of disability, gender, parental education, migration background) and classroom characteristics (average score and variability in the class, class size, teachers with permanent contract). Differences between the treatment and the control groups were small for all variables and not statistically significant.

Compliance

All classes in the treatment group were exposed to the intervention and all classes in the control group were not exposed to the intervention. The large majority of the children actually attended the majority of the lab sessions (98.6% attended at least 3 out of 5 sessions).

[7] **Contini D., Di Tommaso M.L., Mendolia S. (2017),** "The gender gap in mathematics achievement: Evidence from Italian data." Economics of Education Review, 58, pp 32-42.

Attrition
About 5% of the children did not sit the post-test: 5.27% overall, 5.14% in the control group, 5.39 in the treatment group. The difference is not statistically significant. A larger share of children were absent the day of the pre-test (it was held in winter time, many children were home with influenza).

Results
The outcome variable is the post-test score. The main results are shown in Table 1. In columns 1-3 we report the raw differences between treated and control groups. Considering all children (column 1) we find a positive effect of treatment (effect size 0.124), but the effect is statistically significant only for girls (column 3, effect size 0.163). In columns 4-6 we also control for the pre-test score, and in columns 7-9 we include school fixed effects. According to the last two specifications, we find no effect for boys and a large and positive effect for girls (0.154).

In Table 2 we report the results for heterogeneous effects, according to prior achievement. We confirm no effect for boys and a positive effect for girls, increasing with pre-test scores. More specifically, better performing girls benefit the most, and the effect of treatment is statistically significant only for girls with at least average prior performance.

Table 1. Main results

Models Variables	ALL (1)	M (2)	F (3)	ALL (4)	M (5)	F (6)	ALL (7)	M (8)	F (9)
Treatment	0.124 (p=.051)	0.089 (p=.339)	0.163 (p=.050)	0.079 (p=.062)	-0.010 (p=.872)	0.166 (p=.005)	0.079 (p=.011)	-0.023 (p=.612)	0.154 (p=.005)
Pretest scores				0.769 (p=.000)	0.794 (p=.000)	0.735 (p=.000)	0.772 (p=.000)	0.790 (p=.000)	0.746 (p=.000)
School fixed eff							yes	yes	yes
R^2	0.004	0.002	0.007	0.596	0.609	0.577	0.614	0.637	0.603
N	987	485	502	891	442	449	891	442	449

NOTES: Test scores in standardized form
Class clustered errors

Table 2. Heterogeneity of the effects, by prior achievement.

Models Variables	M	F	M	F
Treatment	-0.011 (p=.872)	0.182 (p=.007)	-0.024 (p=.611)	0.168 (p=.002)
Pretest scores	0.790 (p=.000)	0.658 (p=.000)	0.786 (p=.000)	0.689 (p=.000)
Treatment*pretest scores	0.009 (p=.861)	0.161 (p=.014)	0.009 (p=.877)	0.118 (p=.068)
School fixed eff			yes	yes
R^2	0.609	0.583	0.637	0.606
N	442	449	442	449

- **Conclusions:**

The key message is that innovative methodologies have the potential to reduce the gender gap in math. While girls clearly benefit from the intervention, boys are not negatively affected.

However, further investigation is needed. Limitations of the study: (i) small scale, the methodology needs to be tested on a larger group of children; (ii) at present, the study lacks external validity; (iii) the intervention is a short-term activity, it is unlikely to give more than short-term results.