

The effect of positive feedback on primary school students' academic self-concept: Gender heterogeneity in a light-touch randomized intervention

Acta Sociologica

2026, Vol. 69(2) 155–179

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00016993241309552

journals.sagepub.com/home/asj**Tamás Keller** 

HUN-REN Centre for Social Sciences, Computational Social Science—Research Center for Educational and Network Studies, Budapest, Hungary

HUN-REN Centre for Economic and Regional Studies, Institute of Economics, Budapest, Hungary

Abstract

This study analyzes the causal effect of positive feedback on students' task-specific math self-concept using data from a randomized field experiment conducted among rural Hungarian primary school students. It examines how academic self-concept (ASC) responds to the smallest possible dose of positive feedback—a single instance—and explores treatment heterogeneity by gender. The results show that all students who received randomized positive performance feedback experienced a statistically significant (albeit small) improvement in task-specific math self-concept. The positive treatment effect was primarily driven by girls, who experienced a large and statistically significant effect—over 50% greater than the non-significant treatment effect observed among boys. However, the difference in treatment effects between girls and boys, as well as the corresponding decrease in the gender gap between treated and controlled students was not statistically significant. Thus, the results suggest that, while a single instance of positive feedback can temporarily boost students' ASC, it is not a panacea for reducing gender inequalities in ASC. Nevertheless, because girls were particularly responsive to positive feedback treatment and boys were not harmed by it, the results suggest that positive feedback interventions may act as a policy lever for improving girls' self-concept if the intensity of the treatment is enhanced.

Keywords

Positive feedback, academic self-concept, gender gap, randomized experiment, causal effect

Corresponding Author:

Tamás Keller, HUN-REN Centre for Social Sciences, Computational Social Science – Research Center for Educational and Network Studies, Budapest, Hungary.

Email: keller.tamas@tk.hun-ren.hu

Introduction

Girls perform at least as well as boys in mathematics (Guiso et al., 2008; Meinck and Brese, 2019; Neuschmidt et al., 2008; Robinson and Lubienski, 2011). However, in math, girls tend to have a more negative self-concept (Goldman and Penner, 2016; Mejía-Rodríguez et al., 2021; Wilkins, 2004), self-assessment (Mann and DiPrete, 2016), and self-evaluation (Exley and Kessler, 2022) than boys.¹

Girls' lower self-perception in mathematics compared to boys is a global phenomenon. For instance, based on data from the Trends in International Mathematics and Science Study (TIMSS)—an assessment of fourth-grade students—Mejía-Rodríguez et al. (2021) found that in 2015, girls' self-concept in mathematics was lower than boys' in 25 out of the 32 examined countries. This significant gender disparity may play a role in the underrepresentation of girls in STEM² fields (OECD, 2019, p. 171), as individuals often choose to invest their efforts in areas where they feel confident and positive (Correll, 2001; Nagy et al., 2006; Oakes, 1990; Ridgeway, 2011; Sax et al., 2015; Seymour, 1995; Vinni-Laakso et al., 2019). Therefore, gender disparity in schoolchildren's academic self-concept (ASC) may have far-reaching consequences (Barone, 2011; Kriesi and Imdorf, 2019), potentially even contributing to the later gender pay gap (Michelmores and Sassler, 2016; Sterling et al., 2020).

Teachers' feedback practices, in particular, can shape gendered perceptions of ability. When boys receive negative feedback, it tends to focus more on behavior and non-intellectual aspects, with teachers often attributing their failures to a lack of motivation and effort. In contrast, negative feedback for girls primarily targets intellectual inadequacies. As teachers perceive girls as motivated and diligent, they rarely attribute girls' failures to a lack of effort, which reinforces the tendency for girls to attribute their failures to a lack of ability (Dweck et al., 1978). This tendency can ultimately contribute to a gender gap in self-concept, favoring boys. Furthermore, teachers tend to dedicate more instructional time to girls in reading and boys in math (Leinhardt et al., 1979), which may also reinforce gendered perceptions of abilities. For these reasons, girls' ASC in math—their perception of their math ability in school (Marsh and Shavelson, 1985; Shavelson et al., 1976)—needs to be developed using easily accessible forms of leverage (DiPrete and Fox-Williams, 2021). Feedback might be one such solution.

Self-determination theory in psychology suggests that people experience intrinsic motivation when they feel a sense of autonomy and competence (Deci, 1998; Deci et al., 1991; Ryan and Deci, 2000). Positive feedback can significantly impact intrinsic motivation by influencing individuals' perceptions of competence and self-determination (Deci et al., 1975, 2017; Ryan and Deci, 2000). Empirical evidence shows that positive feedback improves achievement. Behncke (2012) describes how students whose teacher read aloud a standard positive affirmation message before their exam scored higher on tests than those who did not receive the positive affirmation. Similarly, highly test-anxious students who read their Facebook friends' affirmation messages before an exam had similar achievements to their peers with low test anxiety (DeLoatch et al., 2017). Furthermore, experimental research in psychology (Katz et al., 2006; Kluger and DeNisi, 1996), economics (Lovász et al., 2022), and sociology (Keller and Szakál, 2021) reveals that positive feedback in the form of encouraging messages boosts students' motivation and increases their persistence. For these reasons, it is important to ask how positive feedback impacts students' ASC.

In observational studies, establishing the causal effect of positive feedback is challenging. Positive feedback typically serves as a reward for good performance, making it difficult to distinguish whether individuals have developed a particular self-concept based on their prior achievements or the feedback they have recently received (Hattie and Timperley, 2007). However, randomized experiments can separate the influence of prior performance from the independent causal effect of feedback through the random allocation of feedback.

To investigate how students' ASC responds to the smallest possible dose of positive feedback—an essential precursor to more extensive interventions—I conducted a randomized field experiment with 1253 Hungarian primary school students from grades 5 to 8, spanning 80 classrooms across 19 schools. The experiment was embedded in a classroom-based, computer-assisted survey. Students first

answered an initial question about their math self-concept and then completed a grade-specific math test. Afterward, a randomly selected half of the students received a single instance of positive automated feedback acknowledging their math performance, regardless of how well they did on the test. The other half of the students received no feedback after the math test. Following this treatment, all students were asked to reassess their math self-concept. This design allowed for the measurement of the short-term causal effect of the light-touch positive feedback treatment on students' self-concept by creating ideal conditions in two key ways. Due to randomization, the feedback was independent of actual performance, allowing for a clear assessment of the treatment's causal effect. Furthermore, repeated self-concept measurement enabled the control of confounding factors related to baseline self-concept.

This study addresses two primary research questions. First, it investigates the causal effect of positive feedback on task-specific math self-concept by analyzing how students' self-concept changes after receiving randomized positive feedback. Second, it examines how the treatment effect varies between girls and boys and, as a consequence of potential treatment heterogeneity, how the initial gender gap in the control group evolves in the treated group.

Concerning the first research question, the preferred specification indicates that the randomized positive feedback treatment led to a 0.15 standard deviation (SD) unit improvement in task-specific math self-concept for students in the treatment group compared to those in the control group. While modest, this effect is noteworthy considering the light-touch nature of the intervention.

Regarding the second research question, the results show that the overall positive treatment effect primarily stemmed from the improvement in girls' task-specific math self-concept. Girls showed a substantially large and statistically significant treatment effect of 0.18 SD units, while boys showed an insignificant treatment effect of about 0.12 SD units. However, the gender difference in treatment effects did not reach statistical significance, suggesting that the treatment might have been too light-touch or the sample size insufficient to detect differences of this magnitude between boys and girls. Therefore, the gender gap in self-concept did not differ significantly between the treated and control groups, although it was smaller and statistically insignificant in the treatment group (0.15 SD units) and larger and statistically significant in the control group (0.2 SD units). In sum, regarding the second research question, the findings are not conclusive. A single instance of positive feedback does not work as a panacea and does not reduce gender inequalities in ASC. However, as girls were particularly responsive to positive feedback and boys were not harmed by it, positive feedback could be an effective lever for improving girls' self-concept in educational practice if the intensity of the one-shot positive feedback is enhanced.

This study is well connected to recent debates and discourses in sociology and economics but is also distinct from prior empirical studies. On the one hand, the research findings indicate that self-related beliefs are malleable to information provision. Thus, the results extend the earlier findings of belief-updating literature (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022) and contribute to the sociological literature on the potential of attitude changes (Bročić and Miles, 2021; Kiley and Vaisey, 2020). Furthermore, the study follows the path of "applied" social research (instead of following the path of "basic" social research) inasmuch as it moves the focus from merely understanding gender inequality to tackling it and promoting social equity (DiPrete and Fox-Williams, 2021).

On the other hand, the study builds on the narrow gender focus of previous experimental research on performance feedback (Lovász et al., 2022; Németh, 1999) but expands the scope by focusing on a younger age group. Unlike prior studies, which examined university students (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022) and young and middle-aged adults (Lovász et al., 2022), this research focuses on primary school students—an age group in which personal traits have not yet crystallized, and self-concept may therefore be more malleable.

The article proceeds as follows: the second section reviews past research, outlines the study's expectations, and establishes its theoretical and empirical context, highlighting the relevance of the study. The third section describes the institutional setting, particularly concerning the gender differences within

Hungarian schools. The fourth section details the experimental design and provides an overview of the sample. The fifth section outlines the empirical strategy used for data analysis. The sixth section presents the results, and the last section concludes with a discussion of the study's limitations and implications.

Review of past research

Positive feedback and self-concept

Positive feedback can take many forms (Hattie and Timperley, 2007), including praise (positive evaluation, see Henderlong and Lepper, 2002), performance feedback (information about the correctness of solutions, see Katz et al., 2006), and encouragement (positive expectations about future performance, see Lovász et al., 2022). Meta-analyses suggest that positive feedback improves performance, but its efficacy hinges on several factors, including the type of feedback and the recipient (Kluger and DeNisi, 1996; Smither et al., 2005). Psychological work further suggests that feedback about the task (or processing of the task) has a greater impact than feedback about the quality of the person (Hattie and Timperley, 2007; Kluger and DeNisi, 1996).

Feedback—both positive and negative—plays a significant role in altering situational self-concept, which is more susceptible to change (Demo, 1992). In this way, interventions that empower students with positive feedback might have the potential to improve situational self-concept (Hattie and Timperley, 2007) such as task-specific self-concept. By contrast, people's general self-concept is known to be constant, as individuals selectively pay attention to feedback that contradicts their initial self-image and reinterpret, diminish, or disregard conflicting feedback (Swann et al., 2003). Therefore, general self-concept may be less malleable through feedback.

For these reasons, I hypothesize that receiving positive performance feedback improves task-specific self-concept (*H1*).

Gender differences in self-concept and the role of positive feedback

There has been extensive sociological research into persistent gender segregation in education, which is highly resistant to change (Barone, 2011; Barone and Assirelli, 2020; Charles and Bradley, 2009; DiPrete and Buchmann, 2013). This body of research highlights the enduring gender gap observed in fields of study such as mathematics and engineering, which have low female representation.

Two main explanations have been proposed to explain gender differences in education: a cultural explanation and a rational choice explanation (Kriesi and Imdorf, 2019). The cultural explanation highlights the influence of internalized gender stereotypes, whereby girls may develop beliefs that they are less capable at math than boys (Correll, 2001; Mann and DiPrete, 2016), while the rational choice explanation posits the influence of the gender-specific cost-benefit considerations behind utility maximization (Jonsson, 1999). Current scholarly discourse tends to lean toward the cultural explanation (Gabay-Egozi et al., 2015; van de Werfhorst, 2017). Recent field experiments have yielded further evidence that aligns with the cultural explanation (Finger et al., 2020). Notably, the cultural explanation for gender differences implies that gender differences may be malleable if girls' negative beliefs about their math abilities can be changed. One possible means of achieving this goal is the provision of feedback (Behncke, 2012; Deloatch et al., 2017; Keller and Szakál, 2021; Lovász et al., 2022).

Females might react more intensively to feedback than males. A small-case ($n = 80$) psychological study at Stanford University indicated that female undergraduate students' self-evaluations were influenced by both positive and negative feedback, while male students were more affected by positive feedback and less impacted by negative feedback (Roberts and Nolen-Hoeksema, 1989). Furthermore, a larger study in the field of economics ($n = 397$) indicated that Eastern European women between the ages of 18 and 45 exhibited greater persistence in a 2-minute online game than their male counterparts when they received regular encouragement messages, such as "You can do it!" (Lovász et al., 2022).

There are multiple reasons why females may have a greater need for positive feedback than males. Studies have shown that women often perform worse than men in competitive settings (Gërxhani et al., 2023), leading them to avoid such situations (Niederle and Vesterlund, 2007; Van Veldhuizen, 2022). This avoidance of competition may indicate a heightened need for empowering positive feedback among females. Furthermore, females are typically more interpersonally sensitive and concerned with others' evaluations (Deci et al., 1975; Katz et al., 2006), which could make them more receptive to feedback. Last, females may have weaker stress-coping abilities (Graves et al., 2021; Matud, 2004), potentially increasing their reliance on positive feedback for support. Therefore, offering positive feedback to females could help promote gender equality in persistence and performance (Lovász et al., 2022), competitiveness (Wozniak et al., 2014), and self-efficacy (Roberts and Nolen-Hoeksema, 1989).

For these reasons, I hypothesize that the effect of positive feedback treatment will be larger for girls and lower for boys (*H2*).

Literature on belief updating

The literature on belief updating in economics examines how individuals adjust their performance beliefs after receiving feedback (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022). This research examines scenarios in which students receive relative performance feedback after completing ability-demanding tasks, such as being informed of their rank within an ability distribution. The central question is how students adjust their initial beliefs in response to this feedback. The feedback they receive can be positive (telling students that their score is higher than a certain percentage of their peers) or negative (telling students that their score is lower than a certain percentage of their peers) and is accurate with a probability of p and inaccurate with a probability of $1 - p$.

These studies show that individuals tend to update their performance beliefs conservatively, deviating much less from their initial beliefs than the Bayesian updating rule would predict. However, there is a notable gender difference, with women demonstrating greater conservatism in belief updating compared to men. The literature is less clear about whether individuals react more strongly to positive or negative feedback. Some studies suggest an asymmetry in responsiveness, with individuals showing greater sensitivity to positive than negative feedback (Eil and Rao, 2011; Möbius et al., 2022), while others find the opposite asymmetry, indicating that people respond more to negative feedback than positive feedback (Coutts, 2019; Ertac, 2011) or even identify little evidence for asymmetry (Buser et al., 2018).

Expanding on the conditions of prior literature

This study expands on the results of previous studies on belief updating in three key aspects: the exclusive focus on relative comparison, the gender difference in the treatment effect, and the effect of feedback on gender inequality.

Relative comparison (Suls et al., 2002) is a widely recognized strategy for evaluating and assessing one's performance. Its role in shaping academic self-concept has been well-documented (Marsh, 1987; Marsh and Parker, 1984). As objective criteria for self-evaluation are often not available, people often rely on comparisons with others to estimate their own outcomes (Festinger, 1954).

However, in many everyday situations, individuals assess their abilities in absolute terms, focusing solely on their own performance without a clear comparative benchmark (Exley and Kessler, 2022; Haaland et al., 2023; Moore and Klein, 2008). Notably, most prior studies on belief updating have relied on relative comparison by providing performance feedback framed in relation to others. As a result, there is limited understanding of how individuals revise their performance beliefs when feedback is given without reference to others' performance (Moore and Klein, 2008).

Another aspect of the literature on belief updating that warrants deeper investigation is the gender difference in the treatment effect. Research has indicated that females tend to update their beliefs less than males, leading to a more conservative approach to updating. However, this evidence of females' rigidity in updating self-related beliefs contrasts with other research suggesting that females are more responsive to positive feedback than males (Lovász et al., 2022; Németh, 1999; Roberts and Nolen-Hoeksema, 1989; Wozniak et al., 2014). Therefore, further research is needed to better understand gender differences in self-concept following positive feedback.

Finally, studies on belief updating have consistently demonstrated that positive feedback enhances self-concept while negative feedback diminishes it. Considering that the initial gender gap favors males' self-concept (Goldman and Penner, 2016; Mejía-Rodríguez et al., 2021; Wilkins, 2004), and assuming that females may respond more strongly to feedback than males (Deci et al., 1975; Katz et al., 2006; Lovász et al., 2022), gender inequality in self-concept could decrease if females improve their self-concept more than males after receiving *positive* feedback. Conversely, gender inequality in self-concept may be exacerbated if females experience a greater decline in their self-concept than males following *negative* feedback. The inequality-exacerbating aspect of negative feedback could have adverse societal implications, as it worsens rather than alleviates existing gender disparities. This potential adverse effect can be avoided by exclusively providing positive feedback.

Setting: gender differences in Hungarian schools

This study explores the impact of positive feedback on students' task-specific math self-concept in rural Hungarian primary schools within a country where gender disparities in the labor market and education favor males, though not significantly more than in other European countries (Horn and Keller, 2015).

In Hungary, primary education lasts for 8 years, starting at age 6 and covering both primary and lower secondary levels—ISCED 1 and ISCED 2—according to the International Standard Classification of Education. In rural Hungarian primary schools, the predominant teaching style is frontal, characterized by students following teachers' instructions and primarily receiving explanations from them. Collaborative activities and group work are comparatively less emphasized in the daily routine of these schools. This educational setup places a high value on teacher feedback; with limited opportunities for peer collaboration, teachers become the primary source from which students receive qualitative assessments of their performance.

Gender differences are evident in Hungarian schools. Females are more altruistic than males but also show lower risk tolerance, lower levels of trust, lower trustworthiness, and lower competitiveness (Horn et al., 2022). Gender disparities in fourth-grade Hungarian students' math performance have escalated over the past two decades, as indicated by TIMSS. Between 2003 and 2015, the gender gap in math performance was statistically insignificant. However, by 2019, it had notably widened, equivalent to 11% SD units in favor of boys (Mullis et al., 2020; Neuschmidt et al., 2008).

Similarly, the gender gap in fourth-grade students' math self-concept has widened, as TIMSS data show. Figure A1 in the Appendix illustrates this trend by depicting the proportion of girls and boys in Hungary who "strongly agree" with the statement "I usually do well in mathematics." Over the past 20 years, the gender gap has almost tripled, from a difference of 5% points ($p = 0.01$) to a difference of 13% points ($p < 0.01$) in favor of boys. Despite this trend, Hungary ranks in the middle among European countries regarding gender differences in math self-concept (Mejía-Rodríguez et al., 2021). In Germany, the Netherlands, and England, the gender gap in math self-concept is twice as large as in Hungary, at around 20% points. Conversely, in Sweden and Cyprus, the gender gap is half as large as in Hungary, at around 6% points (see Figure A2 in the Appendix).

Furthermore, according to data from the Program for International Student Assessment (PISA), considering 15-year-old students who performed best in mathematics, girls are about 10% points less likely to have a career in science and engineering than boys. Despite this notable gender gap, Hungary remains positioned near the middle among European countries (see Figure A3 in the Appendix).

In summary, Hungary is an example of a country where the gender gap in mathematics, both in terms of achievement and self-concept, is widening but remains moderate compared to the broader European context.

Study design and sample description

Sample

Participating schools in this study had previously been contacted for my prior field experiments. I recruited schools by contacting all primary schools in seven contiguous counties of central Hungary in 2017 and used the data to conduct a field experiment in 2018 (Keller and Elwert, 2023). I obtained initial participation agreements from 55 schools. I then refreshed the initial sample to conduct another field experiment in 2020 (Keller, 2020). Out of the schools in the 2018 experiment, 13 agreed to join the new study, and 16 additional schools were newly recruited in 2020, resulting in 29 schools in the second field experiment.

The sample used in this study comes from the voluntary follow-up survey of the 2020 experiment. Out of the 29 schools, 19 schools participated in the recent survey, representing 80 classrooms. Among the participating classrooms, the median classroom size was 16 students, with the maximum and minimum classroom sizes being 24 and 7 students.³

Students' participation in the survey was contingent upon written parental consent obtained from previous experiments and new consent regarding participation in the recent follow-up survey. Thus, students' non-participation was due to the lack of parental consent or random absenteeism from school on the survey day. Based on verbal communication from teachers, most students in the involved classrooms participated in the survey.

The 19 participating schools in the sample are not representative of Hungarian primary schools. School-level comparison based on administrative data suggests that the participating schools are more likely to be small-sized rural schools with below-average performing students than non-participating schools. As Table A1 in the Appendix shows, there are notable differences between participating and non-participating schools. These differences can be substantial, with disparities reaching up to half a standard deviation in math test scores, reading test scores, and students' socioeconomic status.

Experiment

The experiment was embedded in a computer-assisted online student survey, which students filled out in the school's computer lab. The survey was conducted between 20 November 2020 and 19 February 2021 and involved 1253 students from grades 5 through 8. Students participated in the survey in school during a regular school day and were supervised by their teachers. They had 45 minutes to complete the survey.

Supplementary materials, data, and all analytical scripts are archived on the project page at the Open Science Framework: <https://osf.io/3ry8b/>. The study underwent ethical review and received approval from the Institutional Review Board at the HUN-REN Center for Social Sciences, Budapest.

Experimental procedure

Figure 1 shows the experimental procedure. First, students answered a baseline question about their a priori task-specific math self-concept. They then solved the grade-specific math test, followed by the

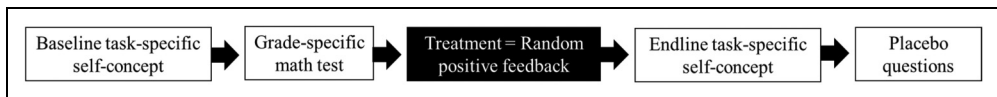


Figure 1. The experimental procedure.

treatment. After the treatment, students evaluated their task-specific math self-concept for the second time. The questionnaire ended with placebo outcome questions to check whether the treatment effect targeted task-specific self-concept (as intended) or had a broader impact on other outcomes.

Treatment. The treatment was integrated into the survey and provided participants with positive feedback. After students had solved the math test, the positive feedback appeared on the computer screen as an automated message. The translated English version of the Hungarian treatment message was: “You did an outstanding job on the math test! As your test score reflects your ability, you should be really proud of yourself as you are a bright and intelligent student.”⁴ Control group students did not receive positive feedback but were directed to proceed to the next question instead.

The treatment combined two types of feedback, in line with the distinction made by Hattie and Timperley (2007). On the one hand, the treatment provided task-specific feedback telling students how well they did on a particular task (“You did an excellent job on the math test”). Task-specific feedback is known to be effective in improving strategies and enhancing self-regulation. On the other hand, the feedback included information about the student as a person (“You are a bright and intelligent student”). Self-specific feedback is considered less effective than task-specific feedback. The two types of feedback were mixed in the treatment message to reflect how teachers typically provide feedback in their daily school routines, as teachers often mix task-specific and self-specific feedback (Airasian, 1997; Bennett and Kell, 1989).

Responding to the limitations of the belief updating literature (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022), the treatment offered absolute criteria for self-evaluation. It also eliminated the potential inequality-exacerbating effect of negative feedback since only positive feedback was provided as the treatment.

Randomization and balance. The randomization of the treatment occurred at the individual level and was based on the value of a randomly generated number. Approximately half of the students were randomly assigned to the treated group, while the other half were assigned to the control group.

Individual-level randomization had two significant implications for the study design. First, it led to heterogeneity within the same classroom, meaning that treated and controlled students could potentially be classmates. This setup provides greater statistical power than randomization at the classroom level. Second, students received the treatment regardless of their actual performance on the test. This approach allows for the establishment of the net treatment effect without the potential contamination of prior performance.

The randomization resulted in a good balance between students assigned to the treated and control groups. As Table A2 in the Appendix summarizes, differences between the treated and control groups were substantively small and statistically not significant at the 5% significance level.

Measurement of variables

Baseline variables. Survey questions asked about students’ academic self-concept (Eccles, 1983; Eccles et al., 1989; Musu-Gillette et al., 2015) and their task-specific math self-concept, followed by a grade-specific math test. Because both the self-concept questions and the math test were administered before the treatment, these measures serve as baseline variables.

Students’ academic self-concept in math was measured by the standardly used and validated survey question by Eccles: “In your opinion, how good are you at math?” (Eccles, 1983; Eccles et al., 1989; Musu-Gillette et al., 2015). Answer categories ranged from 1 (“I am very bad at math”) via 4 (“I am average at math”) to 7 (“I am very good at math”).

Students assessed their task-specific math self-concept twice in the questionnaire: once before the grade-specific math test and again after the treatment. The initial survey question was as follows: “You will shortly solve a short math test. Before you start, please let us know how good you are at

math tests.” To respond, students used a scale from 0 to 10, where 0 indicated “I am not good at all” and 10 indicated “I am excellent.” Students could choose any number between 0 and 10 to express their opinion more accurately.⁵

Student’s baseline task-specific math self-concept exhibited a significant correlation with the commonly used Eccles question for academic math self-concept, with a correlation coefficient of 0.8 ($p < 0.01$). This high correlation coefficient suggests that students’ general math self-concept is highly related to their task-specific (math-test-related) self-concept.

The grade-specific math test used in this study was developed by the Hungarian Educational Authority, drawing on questions from the test banks of the PISA-like National Assessment of Basic Competencies. The test consisted of six problems where students had to employ their math knowledge to solve practical exercises. Test scores refer to the percentage of the correct answers, so the variable ranges between “0” and “1.”

The homeroom teacher reported other baseline variables. Students’ math grade refers to the end-of-term school mark from the second (Spring) semester of the 2019/20 academic year, the school

Table 1. Descriptive statistics of baseline variables.

Stats	Girl	Age	Math grades	Math test scores	ASC in math (Eccles-question)	A priori task-specific math self-concept	Treated
All students							
Mean	0.45	12.95	3.49	0.51	4.45	5.29	0.51
Standard Deviation (SD)	0.50	1.21	1.04	0.25	1.63	2.57	0.50
Percentage of missing	0.00	0.48	7.90	0.00	1.36	3.11	0.00
Observations	1253	1247	1154	1253	1236	1214	1253
Only girls							
Mean	1	12.91	3.57	0.51	4.42	5.15	0.53
Standard Deviation (SD)	0	1.20	1.04	0.25	1.60	2.54	0.50
Percentage of missing	0.00	0.35	6.91	0.00	1.42	3.37	0.00
Observations	564	562	525	564	556	545	564
Only boys							
Mean	0	12.99	3.43	0.52	4.47	5.40	0.49
Standard Deviation (SD)	0	1.21	1.05	0.24	1.65	2.59	0.50
Percentage of missing	0.00	0.58	8.71	0.00	1.31	2.90	0.00
Observations	689	685	629	689	680	669	689

ASC: academic self-concept.

The variables are defined as follows:

Girl: A dummy variable (0/1), where 1 refers to girls and 0 refers to boys.

Age: the difference between the date of the actual survey and the student’s birthday divided by 365. This ranges from 9.55 to 15.84 years.

Math grades are integers from 1 to 5, with 1 being unsatisfactory and 5 being excellent. The scale is as follows: 5 (excellent), 4 (good), 3 (average), 2 (satisfactory), and 1 (unsatisfactory).

Math test scores: Range between 0 and 1, based on the percentage of correct answers in the test.

ASC in math (Eccles-question): Scaled from 1 to 7, where 1 means “I am very bad at math,” 4 means “I am average at math,” and 7 means “I am very good at math.”

A priori task-specific math self-concept: Scaled from 0 to 10, where 0 means “I am not good at all,” and 10 means “I am excellent.”

Treated: A dummy variable (0/1), where 1 indicates treated students.

year before the experiment. Math grades are assigned as integers ranging from 1 to 5. The grading scale is defined as follows: excellent (5), good (4), average (3), satisfactory (2), and unsatisfactory (1). Teachers also reported students' binary gender.

Descriptive statistics about the baseline variables are summarized in Table 1. Participating students were 12.95 years old on average ($SD = 1.21$); 45% were girls and had a medium-level math performance. Students' average achievement on the grade-specific math test scores was 51%, and students' average teacher-awarded math grades were 3.5 on a scale of 1 to 5. The average student estimated themselves as slightly "above average" as their ASC in math (Eccles-question) was 4.45 ($SD = 1.63$) using the seven-grade scale with the theoretical midpoint of the scale at 4. Similarly, the average students evaluated the task-specific math performance as also above average and scored 5.29 on a scale ranging between 0 and 10, with a theoretical midpoint of the scale at 5. These figures indicate the well-documented above-the-average effect (Kruger and Dunning, 1999).

Figure 2 presents descriptive statistics about the initial gender gap in baseline task-specific math self-concept. The left panel of Figure 2 shows that girls outperformed boys in teacher-awarded math grades by 0.14 SD units ($p = 0.052$). However, their performance on math tests did not differ significantly from boys, with a standardized mean difference of -0.05 SD units ($p = 0.25$).

The right panel of Figure 2 shows the gender difference after adjusting for math grades. The negative gap means that, on average, boys score higher than girls. The gender difference in math ASC (Eccles-question) is 11% of the SD ($p = 0.03$). An even larger gender gap can be seen in students' task-

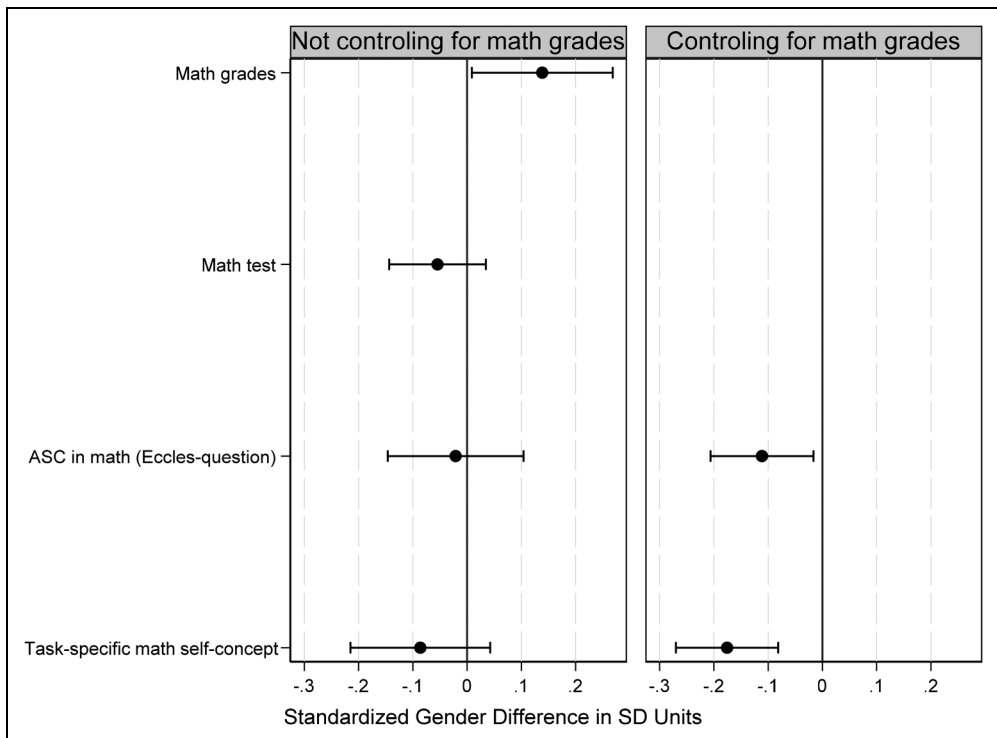


Figure 2. Initial gender gap in math performance and self-concept.

Note: The full results of the regression are contained in Appendix Table A3. A positive gap indicates that girls outperform boys, while a negative gap means that, on average, boys score higher than girls. ASC: academic self-concept.

Table 2. Descriptive statistics of the outcome variables assessed after the treatment.

	Self-concept	Happiness	Inspiration	Acknowledgment	Respect
	How good are you at math tests?	How happy do you feel?	How inspired do you feel?	How much do you feel that people acknowledge you?	How much do you feel that people respect you?
All students					
Mean	5.37	6.85	5.74	5.65	5.56
Standard Deviation (SD)	2.61	2.92	2.91	2.92	3.06
Percentage of missing	0.00	-2.47	-4.23	-6.07	-5.03
Observations	1253	1222	1200	1177	1190
Only girls					
Mean	5.20	6.51	5.47	5.28	5.24
Standard Deviation (SD)	2.60	3.11	3.03	3.00	3.06
Percentage of missing	0.00	-1.95	-3.37	-5.14	-3.72
Observations	564	553	545	535	543
Only boys					
Mean	5.50	7.12	5.97	5.95	5.82
Standard Deviation (SD)	2.61	2.73	2.79	2.81	3.03
Percentage of missing	0.00	-2.90	-4.93	-6.82	-6.10
Observations	689	669	655	642	647

Responses to each variable could be given on a scale ranging from 0 to 10. Endline task-specific math self-concept: Scaled from 0 to 10, where 0 means "I am not good at all," and 10 means "I am excellent." Concerning the other outcome variables, 0 means "Not at all," and 10 means "To a very great extent."

specific math self-concept, with girls providing 0.18 SD units lower estimations than boys ($p = 0.01$). These significant gender differences in self-concept highlight the need for strategies to enhance girls' self-concept, especially given their superior performance in mathematics compared to boys, as evidenced by their math grades.

Outcome variables. Following the treatment, all students completed a set of outcome questions using a 0–10 scale. Initially, they responded to the same task-specific self-concept question for the second time: "Please let us know how good you are at math tests?" This question is referred to as the endline task-specific math self-concept to distinguish it from the baseline task-specific math self-concept asked before the math test and treatment.

Additionally, all students answered placebo outcome questions regarding their current mood, including the following questions: "How happy do you feel?," "How inspired do you feel?," "How much do you feel that people acknowledge you?," and "How much do you feel that people respect you?." These questions were posed only once, as they had not been asked prior to the treatment. The purpose of including these placebo questions was to discern whether the treatment effect targeted task-specific self-concept (as intended) or had a broader, less specific impact.

Descriptive statistics concerning the outcome variables are provided in Table 2.

Empirical strategy

To estimate the treatment effect and test H1, I used equations (1) and (2). Both are classroom fixed-effect ordinary least squares (OLS) linear regression models. Fixed-effects regressions are preferred to control for unobserved heterogeneity at the classroom level in the form of teacher effects that could bias the

results. For example, classrooms might differ in the positive feedback they received initially from the teacher, which could lead to differences between classrooms in how students respond to the treatment. Since the classroom fixed-effect regression compares treated and controlled students within the same classrooms, all unobserved differences between classrooms are controlled for.

In equation (1), the variable $Y_{i,c}$ denotes the i -th student endline task-specific math self-concept in classroom c . The variable $T_{i,c}$ shows whether student i was randomly assigned to the treated ($T=1$) or control group ($T=0$). The variable $F_{i,c}$ signals whether the student is a girl ($F=1$) or a boy ($F=0$). The vector $X_{i,c}$ deploys the control variables. In the preferred specification of equation (1), $X_{i,c}$ contains only students' baseline teacher-awarded grades in math. In other specifications of equation (1), $X_{i,c}$ includes not only students' baseline teacher-awarded grades in math but also math test scores and baseline ASC in math, as measured by the standard Eccles question (Eccles et al., 1983). The variable δ_c stands for classroom fixed effects and $\varepsilon_{i,c}$ is the individual error term clustered at the school level.

The coefficient of interest is β_1 , which shows the causal effect of the treatment on students' task-specific math self-concept. As the treatment was randomized, this coefficient has a causal interpretation. None of the other coefficients in the model have a causal interpretation, as these variables were not randomized.

$$Y_{i,c} = \alpha + \beta_1 \times T_{i,c} + \beta_2 \times F_{i,c} + \beta_3 \times X_{i,c} + \delta_c + \varepsilon_{i,c} \quad (1)$$

The difference between equations (1) and (2) is that equation (2) controls for students' baseline task-specific math self-concept: $Y - 1_{i,c}$. Since the treatment was randomized, including $Y - 1_{i,c}$ into equation (2) does not change the size and interpretation of the treatment effect (β_1). However, it does change the interpretation of the gender gap (β_2), as, in contrast to the randomized treatment, students' gender correlates with the control variables. In equation (2), the gender gap refers to the newly emerged gender difference that arose after students solved the math test since the initial differences in students' task-specific math self-concept are controlled for by $Y - 1_{i,c}$.

$$Y_{i,c} = \alpha + \beta_1 \times T_{i,c} + \beta_2 \times F_{i,c} + \beta_3 \times X_{i,c} + \beta_4 \times (Y - 1_{i,c}) + \delta_c + \varepsilon_{s,c} \quad (2)$$

To estimate the gender heterogeneity in the treatment and test H2, I used equations (3) and (4). Both of these models contain the two-way interaction between the treatment and students' gender ($T_{i,c} \times F_{i,c}$).

In estimating gender heterogeneity in the treatment, the parameters of interest are the coefficients γ_1 and $\gamma_1 + \gamma_3$. These coefficients show the treatment effect for boys (γ_1) and girls ($\gamma_1 + \gamma_3$). Another parameter of interest is the gender gap in the control (γ_2) and treated groups ($\gamma_2 + \gamma_3$). The coefficient γ_3 has no causal interpretation, as students' gender cannot be randomized in a similar way to how the treatment was randomized. The coefficient γ_3 tests whether the treatment effect is statistically different between boys and girls.

$$Y_{i,c} = \alpha + \gamma_1 \times T_{i,c} + \gamma_2 \times F_{i,c} + \gamma_3 \times (T_{i,c} \times F_{i,c}) + \gamma_4 \times X_{i,c} + \delta_c + \varepsilon_{i,c} \quad (3)$$

Since equation (4) controls for students' a priori task-specific math self-concept, the gender gaps in the control and treated groups (γ_2 and $\gamma_2 + \gamma_3$) refer to the gender gaps that emerged after students solved the math test. However, including students' baseline self-concept does not affect the treatment effect among boys and girls (γ_1 and $\gamma_1 + \gamma_3$) as the treatment was randomized and did not correlate with any covariates.

$$Y_{i,c} = \alpha + \gamma_1 \times T_{i,c} + \gamma_2 \times F_{i,c} + \gamma_3 \times (T_{i,c} \times F_{i,c}) + \gamma_4 \times X_{i,c} + \gamma_5 \times (Y - 1_{i,c}) + \delta_c + \varepsilon_{i,c} \quad (4)$$

Results

Descriptive results

Figure 3 shows the raw mean difference in task-specific math self-concept, calculated as the change from baseline to endline. The data is categorized by treatment status (control/treated) and gender (boys/girls).

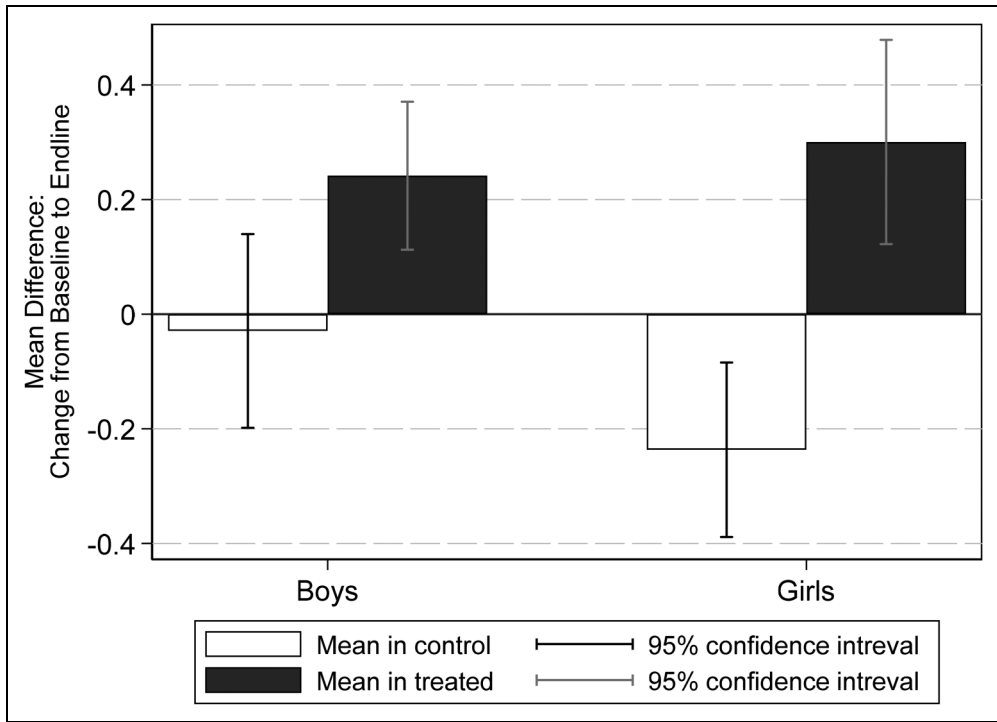


Figure 3. Change in task-specific math self-concept from baseline to endline, categorized by treatment status (control/treated) and gender (boys/girls).

In the treatment group (represented by black bars), students experienced a positive update in their self-concept. In the control group (represented by white bars), where students did not receive positive feedback, both girls' and boys' self-concept decreased after completing the math test. The negative shift in the control group suggests that engaging in a demanding task can reduce self-concept. Such negative updates underscore the importance of a surplus in self-concept when undertaking challenging tasks. The decrease in self-concept after completing ability-demanding tasks also emphasizes the importance of replenishing reduced self-concept through positive feedback.

The figure also suggests the treatment effect, which is represented by the difference between the white and black bars. Girls showed a more pronounced treatment effect compared to boys, as girls in the control group experienced a sizable decline in their task-specific math self-concept after completing the math test (leading to a larger difference between the white and black bars). In contrast, boys in the control group showed a smaller decline in their task-specific self-concept (thus, the difference between the white and black bars is smaller for boys than for girls).

Main treatment effect: the test of H1

Table 3 presents the estimations using both equations (1) and (2). The intervention resulted in a significant improvement in students' task-specific math self-concept. The treatment effect is equal to a 0.3–0.4 unit improvement on the natural scale of the dependent variable, where the average of students' task-specific math self-concept is 5.25. Expressing this treatment effect in SD units yields a small effect size between 0.12 and 0.15 SD units. Thus, the results provide causal support for H1, which posits that receiving positive feedback increases students' task-specific math self-concept.

Table 3. Treatment effect on endline task-specific math self-concept—unstandardized OLS regression coefficients.

Estimated by	Equation (1)			Equation (2)	
	(1)	(2)	(3)	(4)	(5)
Treated (β_1)	0.32*	0.39**	0.37*	0.37**	0.37**
	(0.14)	(0.13)	(0.13)	(0.08)	(0.09)
Girls (β_2)		-0.49**	-0.45**	-0.13	-0.13
		(0.14)	(0.15)	(0.09)	(0.09)
Math grade = 1 (worst)		-3.99**	-3.59**	-0.65	-0.32
		(0.60)	(0.69)	(0.56)	(0.63)
Math grade = 2		-5.58**	-5.04**	-0.53+	-0.18
		(0.38)	(0.46)	(0.30)	(0.28)
Math grade = 3		-4.59**	-4.14**	-0.46+	-0.17
		(0.29)	(0.33)	(0.24)	(0.21)
Math grade = 4		-3.67**	-3.30**	-0.31	-0.10
		(0.24)	(0.29)	(0.20)	(0.17)
Math grade = 5 (best)				Ref.	Ref.
Math test score			1.85**	0.68*	0.65*
			(0.40)	(0.26)	(0.25)
ASC in math (Eccles-question)					0.21*
					(0.08)
A priori task-specific math self-concept				0.80**	0.72**
				(0.03)	(0.06)
Constant	5.20**	8.09**	6.84**	0.90*	0.24
	(0.07)	(0.19)	(0.41)	(0.38)	(0.30)
Observations	1253	1253	1253	1253	1253
R-squared	0.13	0.41	0.43	0.74	0.75
Treatment effect in SD units	0.12	0.15	0.14	0.14	0.14
Mean in the control group	5.25	5.25	5.25	5.25	5.25

Robust standard errors (clustered at the school level) are in parentheses. ASC: academic self-concept.

** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

All models include classroom fixed effects. Missing values in the baseline variables are replaced by zero, and a separate dummy variable controls for missing status (these variables are not included in the table).

Treatment heterogeneity by gender: the test of H2

Figure 4 visualizes the heterogeneity in the treatment effect using both equation (3) (point estimates depicted with black circles) and equation (4) (point estimates depicted with gray diamonds), with the full regression models presented in Table A4 in the Appendix. Unlike the main treatment effect, which is causal, the analysis of treatment heterogeneity is exploratory and not causal, as students' gender could not be randomized in the same way as the treatment.

The treatment effect appears statistically insignificant for boys ($\gamma_1 = 0.31$; expressed in SD units = 0.12; $p = 0.16$). However, a statistically significant and positive treatment effect is observed among girls. Treated girls assessed their task-specific math self-concept as $\gamma_1 + \gamma_3 = 0.48$ units higher ($p = 0.02$) than girls in the control group, equating to a treatment effect of 0.18 SD units. Thus, the treatment effect is over 50% higher among girls than among boys. However, as Figure 4 illustrates, the gender-specific treatment effects are imprecisely estimated with large standard errors (SE), resulting in a statistically insignificant difference between the treatment effect among boys and girls. ($\gamma_3 = 0.17$; $p = 0.57$).

A similar pattern is observed in gender-specific treatment effect when controlling for students' a priori task-specific math self-concept (equation (4)). The treatment effect primarily concentrates on girls, yielding a statistically significant treatment effect among them ($\gamma_1 + \gamma_3 = 0.50$; expressed in SD units = 0.19;

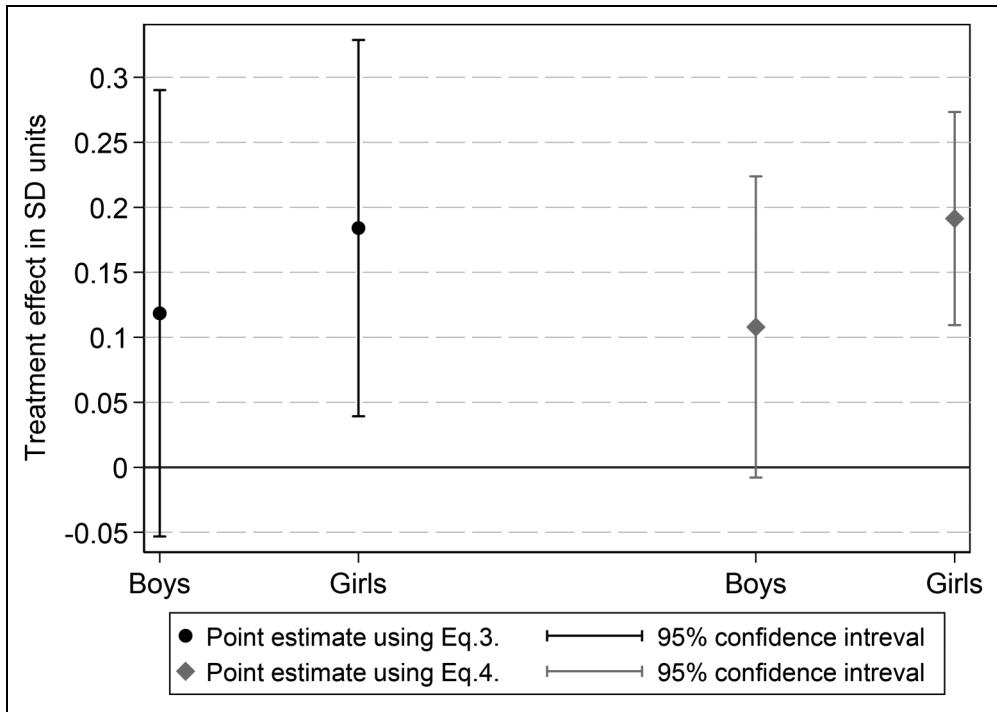


Figure 4. Treatment effect by gender.

$p < 0.01$), but the treatment effect is insignificant at the 5% level for boys ($\gamma_1 = 0.28$; expressed in SD units = 0.11; $p = 0.07$). The gender difference in the treatment effect is not statistically significant ($\gamma_3 = 0.22$; $p = 0.26$). In summary, the results do not support H2, as the difference between the treatment effects for boys and girls did not reach statistical significance.⁶

Related to the insignificant gender difference in the treatment effect, the gender gap also does not differ between treated and control groups. Relying on the results estimated by equation (3), the initial gender gap in the control group ($\gamma_2 = 0.58$, equating to a gender gap of -0.2 SD units; $p = 0.01$) is reduced by 40% in the treatment group ($\gamma_2 + \gamma_3 = -0.4$, corresponding to a gender gap of -0.15 SD units; $p = 0.07$). However, the difference between the gender gaps before and after the treatment is statistically not significant ($\gamma_3 = 0.17$; $p = 0.57$). See Figure A4 in the Appendix.

Robustness checks

Several robustness tests were conducted to validate the results. Figure 5 presents a sensitivity analysis of treatment heterogeneity based on students' initial math performance (as measured by math grades and recent test scores) and academic self-concept (measured by the Eccles question and task-specific math self-concept). The findings indicate that the positive feedback had a substantially larger impact on students with higher initial math grades and stronger ASCs. However, the treatment also had a positive impact on students with average or low math test scores and self-concept, though it did not influence those with average or low math grades.

Treatment heterogeneity based on students' recent math test scores is particularly relevant. Students with high test scores received (honest) positive feedback as a reward for their excellent performance, while those with average or low scores received biased positive feedback that inaccurately acknowledged

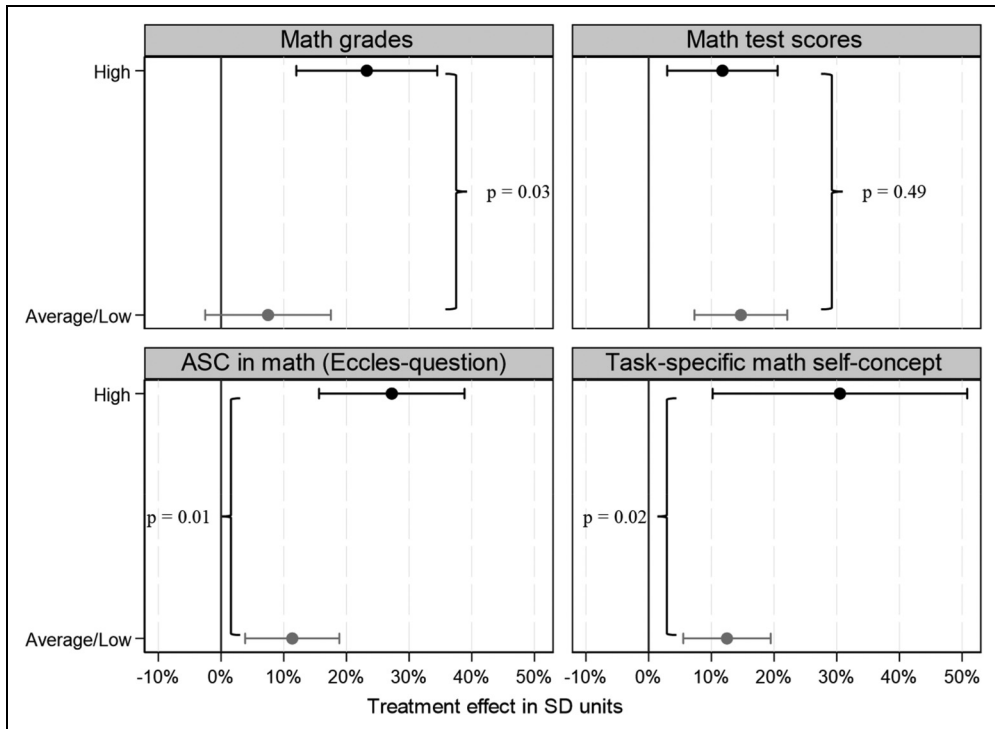


Figure 5. Treatment heterogeneity according to (level of) math performance and baseline math self-concept.

ASC: academic self-concept.

The figure shows various treatment effects estimated by separate classroom fixed-effect OLS regression results estimated by equation (2) and including only prior math grades as a control variable into $X_{i,c}$. Full regression results are in Table A5 in the Appendix.

The separate OLS regressions were estimated in the “high” and “average/low” categories of the corresponding baseline variable, defined as follows:

Average/low math grades: Students with a math grade of 3 or lower in the last semester’s report card.

High math grades: Students with a math grade of 4 or 5 in the last semester’s report card.

Average/low math test scores: Students with a 70% or lower performance on the math test.

High math test scores: Students with higher than 70% performance on the math test.

Average/low ASC in math (Eccles-question): Students who said they were “very bad,” “bad,” “a little worse than average,” “average,” or “a little better than average” at math.

High ASC in math (Eccles-question): Students who said they were “good” or “very good” at math.

Average/low task-specific math self-concept: Students who scored 0 through 8 on the test-specific math ASC ranging from 0, “I am not good at all,” to 10, “I am excellent.”

High task-specific math self-concept: Students who scored 9 or 10 on the test-specific math ASC ranging from 0, “I am not good at all,” to 10, “I am excellent.”

Differences between the two treatment effects in each panel were estimated using seemingly unrelated estimations (*suest*), a post-estimation command built into Stata®.

their performance as excellent. Notably, no treatment heterogeneity was observed based on recent math test performance. Students scoring above 70% experienced similar positive changes in their task-specific self-concept as those scoring below 70%. This suggests that the honesty of positive feedback does not generate treatment heterogeneity.

Treatment heterogeneity based on prior math performance was further analyzed with a focus on gender differences, comparing boys and girls with either high or average/low math performance. When defining math performance according to prior math grades, the treatment effect did not differ significantly between boys and girls, regardless of whether their prior grades were average/low or high (see the left panel of Figure A5 in the Appendix). However, when math performance was defined by recent test results, girls who performed well on the math test showed a greater treatment effect than similarly achieving boys (see the right panel of Figure A5 in the Appendix). This suggests that girls' self-concept is more responsive to honest, positive feedback about their recent performance than that of boys.

Further, robustness checks suggest that the treatment exerted its effect as intended and targeted only students' task-specific math self-concept. The positive feedback intervention did not affect students' feelings of happiness, inspiration, acknowledgment, or respect, as the treatment effect was statistically insignificant and substantially small across all models (Table A6 in the Appendix).⁷ Although there were notable gender differences in each of these variables, with girls indicating lower values than boys, the treatment effect did not substantially differ between boys and girls.

Discussion and conclusion

There has been extensive research into the gender gap in education (Barone, 2011; Barone and Assirelli, 2020; Charles and Bradley, 2009; DiPrete and Buchmann, 2013; Kriesi and Imdorf, 2019), though less attention has been paid to the potential tools for mitigating this gap (DiPrete and Fox-Williams, 2021; Legewie and DiPrete, 2012). This study focused on girls' biased self-concepts that may hinder their engagement in math-intensive educational choices (Correll, 2001; Nagy et al., 2006; Oakes, 1990; Ridgeway, 2011; Sax et al., 2015; Seymour, 1995; Vinni-Laakso et al., 2019) but which can be changed through feedback (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022). The study analyzed how positive performance feedback (as a scalable tool in education) affects students' task-specific math self-concept and how it mitigates the initial gender gap in self-concept (which typically favors boys).

A randomized field experiment was conducted with rural Hungarian primary school students to test how their task-specific math self-concept responds to a single instance of randomized positive feedback, representing the smallest possible dose of positive feedback intervention. After answering a baseline question about their task-specific math self-concept and completing a short grade-specific competency-based math test, students received either positive absolute performance feedback or no feedback, determined by a random algorithm and independent of actual performance on the math test. Following the treatment, students were asked about their task-specific math self-concept for a second time.

The design extends prior literature on belief updating in several key ways (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022). First, it targeted primary school-aged students with positive feedback as their academic self-concepts (ASC) may be more malleable than university students' ASC; prior experiments have primarily focused on university students with better-established, less malleable ASCs. Second, students received absolute performance feedback rather than relative feedback. This means that instead of making relative comparisons to peers, students evaluated their ability in general, which aligns more closely with everyday settings (Moore and Klein, 2008). Third, unlike prior research that provided indications about the likelihood of feedback being true or false (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022), students in this study received feedback without any indication of its accuracy.⁸ This approach was chosen to reflect better real-world settings, where individuals often receive feedback without being informed of its truthfulness and must judge its validity themselves. Last, students received only positive feedback to avoid any potential decline in self-concept, which could exacerbate the gender gap.⁹

The results showed that all treated students experienced a positive improvement in their self-concept compared to those in the control group, who did not receive positive feedback. Since the treatment was randomized, the treatment effect is causal. The treatment effect was particularly pronounced for girls,

with a significant positive effect that was 50% higher than the treatment effect for boys. The positive treatment effect was statistically not significant for boys. However, the non-causal gender difference in the treatment effect and the associated reduction in the gender gap between treated and controlled students were statistically *not* significant. Therefore, while a single instance of positive feedback improves students' ASC, it does not reduce the gender gap in self-concept. Nevertheless, because girls' self-concept was particularly responsive to positive feedback, while boys' self-concept was not harmed by it, the results suggest that increasing the frequency of this one-shot positive feedback could serve as a policy lever to improve girls' self-concept.

The results might have implications for status inequality in self-concept. Given that low-status students tend to make lower assessments of their abilities and high-status students tend to make higher ones (Sullivan, 2006), providing targeted positive feedback exclusively to low-status students might help mitigate existing status inequality in self-concept. However, status inequality in self-concept might increase if all students receive positive feedback.

Future research could employ two major strategies to elaborate on the insignificant results concerning the estimation of gender differences in the treatment effect. First, the intensity of the positive feedback could be enhanced by increasing the frequency of the one-shot positive feedback. Delivering feedback orally, especially from a socially significant individual, may enhance its impact compared to the automated written feedback used in this study. Second, the statistical power of the sample could be enhanced by increasing the size of the sample, which consisted of over 1200 students in this study. While this sample size is not small, it may not be sufficiently large to identify the treatment heterogeneity of this magnitude.

In terms of study design, some specific issues warrant discussion. Concerns may arise that self-concept enhancement is not always beneficial. The positive effect of empowerment could lead to inflated self-concepts that are not grounded in reality and could contribute to overconfidence (Kruger and Dunning, 1999; Moore and Healy, 2008).

In contrast, prior research has suggested that positive self-perceptions serve important purposes, such as assisting in the pursuit of goals, influencing others, and tackling complex and challenging tasks (Schwardmann and van der Weele, 2019). This study further expanded on this reasoning by suggesting that individuals might need to have a certain level of overconfidence to engage in tasks that require their abilities. This is because engagement in ability-demanding tasks can potentially lower students' self-concept, as observed in the control group (Figure 3). Consequently, some students may refrain from challenging themselves in demanding situations due to a lack of positive self-concept (Epstein, 1973).¹⁰ This might be particularly evident in educational choices such as applying to knowledge-intensive educational programs or pursuing STEM fields where the possibility of rejection/failure is high (which would send negative signals about students' abilities and decrease self-concept). To engage in these risky choices, students may need a degree of overconfidence.

A related critique concerns the relevance of feedback that was not directly related to students' actual performance and, at least for some students, provided false information. Consequently, the result might also be interpreted as girls being particularly susceptible to disinformation induced by biased feedback.

However, girls were not more susceptible to potential misinformation than boys as the treatment effect does not differ between high/low performing boys and girls if performance is defined based on students' prior math grades (see the left panel of Figure A5 in the Appendix for a reference). Indeed, girls were more receptive to honest and positive feedback about their recent math performance than boys (see the right panel of Figure A5 in the Appendix). Consequently, girls can translate positive feedback about their recent performance into a larger self-concept improvement than boys—an important message that educational practitioners should consider.

A further potential concern is that providing students with positive signals about their abilities without a direct connection to actual performance is artificial. This argument can be further expanded to raise

questions about the study's external validity and whether schools would actually implement this type of feedback.

However, students shape their understanding of their abilities through the feedback they receive, making the development of self-concept an ongoing learning process. Throughout this process, students internalize feedback and cultivate a positive self-image in areas where they have been recognized and acknowledged by their environment (Hattie and Timperley, 2007). Research has also highlighted how social status differences in parenting styles can contribute to the achievement gap observed among students (Bradley and Corwyn, 2002; Kalil and Ryan, 2020). Differences in children's empowerment may be one aspect of differences in parenting styles (Gunderson et al., 2013; Hoff et al., 2002). Furthermore, pedagogical research has underscored the importance of praising students, even for small accomplishments that can be observed among all students regardless of their overall academic performance (Burnett, 2002; Floress and Jenkins, 2015). As a result, the treatment is not far removed from real-life practices since parents (and sometimes also teachers) deliberately create opportunities to praise and support students.

Last, concerns could be raised about the small effect size, and skepticism expressed about the design, which rendered the outcome of the intervention immediate, both temporally (the two self-concept questions were separated by math test and the treatment) and in content (the same question was used in the baseline and endline task-specific self-concept questions). Consequently, it may be posited that it is indeed surprising that contrary to the favorable experimental design, only small treatment effects were found.

However, the modest treatment effect should be interpreted in light of the anchoring effect, which describes people's tendency to maintain consistency in their responses (Furnham and Boo, 2011; Tversky and Kahneman, 1974). Therefore, in this study, students might have recalled their initial answers and provided a similar response to the second task-specific self-concept question. Thus, a significant shock may be required to disrupt the status quo and produce a substantial improvement in self-concept. Furthermore, the favorable experimental design employed in this study is deliberately crafted to maximize the chance of detecting a treatment effect, laying the foundation for more targeted future research that might be able to investigate less proximate outcomes, perhaps even outcomes associated with math achievement.

In conclusion, this study showed that even a single instance of positive feedback can *temporarily* improve students' self-concept. However, this approach cannot work as a panacea to reduce gender inequalities. Nevertheless, as girls were particularly responsive to positive feedback while boys were not harmed by it, positive feedback has the potential to serve as an effective policy lever for enhancing girls' self-concept if the frequency of the one-shot treatment is increased. Therefore, more intensive positive-feedback treatments may hold promise for mitigating or closing the gender gap in educational decisions related to math, such as selecting classes with specialized math courses or pursuing STEM fields at the college level—decisions that, to some extent, all rely on math self-concept (Correll, 2001; Nagy et al., 2006; Oakes, 1990; Ridgeway, 2011; Sax et al., 2015; Seymour, 1995; Vinni-Laakso et al., 2019).


Acknowledgment

The author thanks Carlo Barone and Nevena Kulic and the audiences at the ISA RC28 Spring Meeting 2024 in Shanghai and the Meeting of the Economics of Education Association in 2023 (AEDE) in Santiago de Compostella.

Funding

The research was supported by grants from the Hungarian National Research, Development and Innovation Office (NKFIH), Grant number K-135766; the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/ 00569/21/9) and the New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund (Grant Number: ÚNKP-23-5-CORVINUS-149). Funding from the Horizon Europe project 'EFFEct' grant (no. 101129146) is also gratefully acknowledged.

ORCID iD

Tamás Keller  <https://orcid.org/0000-0001-6943-0955>

Supplemental Material

Supplemental material for this article is available online and on the project page at the Open Science Framework: <https://osf.io/3ry8b/>, where data and all analysis scripts are archived.

Notes

1. Studies use different terminology to describe the phenomenon of individuals reflecting on their abilities. This study adopts the concept of self-concept, one of the broadest and most well-established concepts in social research (Shavelson et al., 1976). In particular, this term is used to study mathematical self-concept in terms of people's perceptions or beliefs about their mathematical abilities (Reyes, 1984).
2. An acronym for science, technology, engineering, and mathematics.
3. There is no precise information regarding the proportion of students who did not participate in the survey from each classroom, as schools only provided data for students who had obtained parental consent. The average classroom size in Hungarian primary education is approximately 20 students per classroom (Varga, 2024). This figure corroborates that most students in the classroom might have participated in the study.
4. The original Hungarian treatment message was: "Fantasztikusan teljesítettél a matek-teszten! Mivel a teszten elért eredményed jól tükrözi a képességeidet igazán büszke lehetsz magadra, hiszen okos és intelligens diák vagy."
5. The original Hungarian survey question was the following: "Mindjárt következnek a matek feladatok. Mielőtt ezekbe belekezdnél, kérek válaszold meg a következő kérdéseket! Mit gondolsz, mennyire vagy jó a matek-tesztekben? Használd az alábbi skálát. A 0 azt jelenti, hogy 'egyáltalán nem vagyok jó,' a 10 pedig azt, hogy 'kiváló vagyok.' Bármely számot választhat 0 és 10 között, hogy véleményed pontosabban kifejezzed."
6. The insignificant gender difference in the treatment effect is not due to the inclusion of classroom fixed effects, which reduce the variance in the dependent variable. Re-estimating the coefficients from Figure 4 without fixed effects yields similar results to those originally presented (the re-estimated figure is not included in the paper).
7. To perform these tests, the dependent variable in equation (1) was changed, and variables listed in Table 2 were used instead of students' task-specific math self-concept. All these variables were assessed once after the treatment.
8. For example, in Coutts (2019), feedback was given by three robots: two always told the truth, while one always lied, but students were unaware of which robot provided their feedback. The robots delivered either positive or negative signals, making true signals occur with a probability of two-thirds. A similar approach was used by Möbius et al. (2022).
9. As negative feedback impacts students' self-concept negatively (Buser et al., 2018; Coutts, 2019; Eil and Rao, 2011; Ertac, 2011; Möbius et al., 2022) and since females typically respond more intensely to feedback (Lovász et al., 2022; Németh, 1999; Roberts and Nolen-Hoeksema, 1989; Wozniak et al., 2014), the gender gap in self-concept could widen in response to negative feedback.
10. However, overconfidence can be detrimental when it leads individuals to engage in risky behaviors, such as driving at high speeds based on an inflated belief in their driving skills or investing in unsafe portfolios due to exaggerated confidence in their financial literacy. However, when it comes to applying to knowledge-intensive educational tracks, pursuing STEM fields, or competing for scholarships, these are not inherently risky behaviors that can harm students' livelihoods.

References

- Airasian PW (1997) *Classroom Assessment*. New York: McGraw-Hill.
- Barone C (2011) Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education* 84(2): 157–176.
- Barone C and Assirelli G (2020) Gender segregation in higher education: An empirical test of seven explanations. *Higher Education* 79(1): 55–78.

- Behncke S (2012) How do shocks to non-cognitive skills affect test scores? *Annals of Economics and Statistics* (107/108): 155.
- Bennett N and Kell J (1989) *A Good Start? Four Year Olds in Infant Schools*. Oxford: Blackwell.
- Bradley RH and Corwyn RF (2002) Socioeconomic Status and child development. *Annual Review of Psychology* 53(1): 371–399.
- Bročić M and Miles A (2021) *College and the “Culture War”: Assessing Higher Education’s Influence on Moral Attitudes*.
- Burnett PC (2002) Teacher praise and feedback and students’ perceptions of the classroom environment. *Educational Psychology* 22(1): 5–16.
- Buser T, Gerhards L and van der Weele J (2018) Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty* 56(2): 165–192.
- Charles M and Bradley K (2009) Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American Journal of Sociology* 114(4): 924–976.
- Correll SJ (2001) Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology* 106(6): 1691–1730.
- Coutts A (2019) Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics* 22(2): 369–395.
- Deci EL (1998) The relation of interest to the motivation of behavior: A self-determination theory perspective. In: Renninger KA, Hidi S and Krapp A (eds) *The Role of Interest in Learning and Development*. Erlbaum: Hillsdale, 43–70.
- Deci EL, Cascio WF and Krusell J (1975) Cognitive evaluation theory and some comments on the Calder and Staw critique. *Journal of Personality and Social Psychology* 31(1): 81–85.
- Deci EL, Olafsen AH and Ryan RM (2017) Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior* 4(1): 19–43.
- Deci EL, Vallerand R, Pelletier L, et al. (1991) Motivation and education: The self-determination perspective. *Educational Psychologist* 26(3): 325–346.
- Deloatch R, Bailey BP, Kirlik A, et al. (2017) I need your encouragement! requesting supportive comments on social media reduces test anxiety. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp.736–747. New York, NY, USA: ACM.
- Demo D (1992) The self-concept over time: Research issues and directions. *Annual Review of Sociology* 18(1): 303–326.
- DiPrete TA and Buchmann C (2013) *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*. New York: Russell Sage Foundation.
- DiPrete TA and Fox-Williams BN (2021) The relevance of inequality research in sociology for inequality reduction. *Socius* 7: 1–30. DOI: 10.1177/23780231211020199.
- Dweck CS, Davidson W, Nelson S, et al. (1978) Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental Psychology* 14(3): 268–276.
- Eccles JS (1983) Expectancies, values, and academic behaviors. In: Spence JT (ed) *Achievement and Achievement Motives: Psychological and Sociological Approaches*. San Francisco, CA: W. H. Freeman, 75–146.
- Eccles JS, Adler TF, Futterman R, et al. (1983) Expectancies, values and academic behaviors. In: Spence JT (eds) *Achievement and Achievement Motives*. San Francisco: W. H. Freeman.
- Eccles JS, Wigfield A, Flanagan CA, et al. (1989) Self-concepts, domain values, and self-esteem: Relations and changes at early adolescence. *Journal of Personality* 57(2): 283–310.

- Eil D and Rao JM (2011) The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics* 3(2): 114–138.
- Epstein S (1973) The self-concept revisited: Or a theory of a theory. *American Psychologist* 28(5): 404–416.
- Ertac S (2011) Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior and Organization* 80(3): 532–545.
- Exley CL and Kessler JB (2022) The gender gap in self-promotion. *The Quarterly Journal of Economics* 137(3): 1345–1381.
- Festinger L (1954) A theory of social comparison processes. *Human Relations* 7(2): 117–140.
- Finger C, Solga H, Ehlert M, et al. (2020) Gender differences in the choice of field of study and the relevance of income information. Insights from a field experiment. *Research in Social Stratification and Mobility* 65: 100457.
- Floress MT and Jenkins LN (2015) A preliminary investigation of kindergarten Teachers' use of praise in general education classrooms. *Preventing School Failure: Alternative Education for Children and Youth* 59(4): 253–262.
- Furnham A and Boo HC (2011) A literature review of the anchoring effect. *The Journal of Socio-Economics* 40(1): 35–42.
- Gabay-Egozi L, Shavit Y and Yaish M (2015) Gender differences in fields of study: The role of significant others and rational choice motivations. *European Sociological Review* 31(3): 284–297.
- Gërxhani K, Brandts J and Schram A (2023) Competition and gender inequality: A comprehensive analysis of effects and mechanisms. *American Journal of Sociology* 129(3): 715–752.
- Goldman AD and Penner AM (2016) Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth* 21(4): 403–418.
- Graves BS, Hall ME, Dias-Karch C, et al. (2021) Gender differences in perceived stress and coping among college students. *PLoS One* 16(8 August): 1–12.
- Gunderson EA, Gripshover SJ, Romero C, et al. (2013) Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development* 84(5): 1526–1541.
- Guiso L, Monte F and Sapienza P (2008) Differences in test scores correlated with indicators of gender equality. *Science* 320(May): 1–2.
- Haaland I, Roth C and Wohlfart J (2023) Designing information provision experiments. *Journal of Economic Literature* 61(1): 3–40.
- Hattie J and Timperley H (2007) The power of feedback. *Review of Educational Research* 77(1): 81–112.
- Henderlong J and Lepper MR (2002) The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin* 128(5): 774–795.
- Hoff E, Laursen B, Tardif T, et al. (2002) Socioeconomic status and parenting. Handbook of parenting Volume 2. *Biology and Ecology of Parenting* 8(2): 231–252.
- Horn D and Keller T (2015) Hungary: The impact of gender culture. In: edited by Blossfeld H-P, Skopek J, Triventi M, et al. (eds) *Gender, Education and Employment*. Cheltenham: Edward Elgar Publishing, 287–303.
- Horn D, Kiss HJ and Lénárd T (2022) Gender differences in preferences of adolescents: Evidence from a large-scale classroom experiment: Gender differences in preferences of adolescents. *Journal of Economic Behavior and Organization* 194: 478–522.
- Jonsson JO (1999) Explaining sex differences in educational choice an empirical assessment of a rational choice model. *European Sociological Review* 15(4): 391–404.
- Kalil A and Ryan R (2020) Parenting practices and socioeconomic gaps in childhood outcomes. *The Future of Children* 30(2020): 29–54.

- Katz I, Assor A, Kanat-Maymon Y, et al. (2006) Interest as a motivational resource: Feedback and gender matter, but interest makes the difference. *Social Psychology of Education* 9(1): 27–42.
- Keller T (2020) Breaking the Negative Spiral of Disruptive School Behavior—Experimental Evidence about How an Exogenous Shock to Primary School Students’ School Behavior Affects Their Achievement. *AEA RCT Registry*, May 13. Initial registration date: February 10, 2020. <https://www.socialscienceregistry.org/trials/5442/history/68070>
- Keller T and Elwert F (2023) Feasible peer effects: Experimental evidence for deskmate effects on educational achievement and inequality. *Sociological Science* 10: 806–829.
- Keller T and Szakál P (2021) Not just words! Effects of a light-touch randomized encouragement intervention on Students’ exam grades, self-efficacy, motivation, and test anxiety. *PLoS One* 16(9): e0256960.
- Kiley K and Vaisey S (2020) Measuring stability and change in personal culture using panel data. *American Sociological Review* 85(3): 477–506.
- Kluger AN and DeNisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119(2): 254–284.
- Kriesi I and Imdorf C (2019) Gender segregation in education. In: Becker R (ed) *Research Handbook on the Sociology of Education*. Cheltenham, UK: Edward Elgar Publishing, 193–212. DOI: 10.4337/9781788110426.00020.
- Kruger J and Dunning D (1999) Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121–1134.
- Legewie J and DiPrete TA (2012) School context and the gender gap in educational achievement. *American Sociological Review* 77(3): 463–485.
- Leinhardt G, Seewald AM and Engel M (1979) Learning what’s taught: Sex differences in instruction. *Journal of Educational Psychology* 71(4): 432–439.
- Lovász A, Cukrowska-Torzewska E, Rigó M, et al. (2022) Gender differences in the effect of subjective feedback in an online game. *Journal of Behavioral and Experimental Economics* 98: 101854.
- Mann A and DiPrete T (2016) The consequences of the national math and science performance environment for gender differences in STEM aspiration. *Sociological Science* 3: 568–603.
- Marsh HW (1987) The Big-Fish-Little-Pond Effect on academic self-concept. *Journal of Educational Psychology* 79(3): 280–295.
- Marsh HW and Parker JW (1984) Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don’t learn to swim as well? *Journal of Personality and Social Psychology* 47(1): 213–231.
- Marsh HW and Shavelson R (1985) Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist* 20(3): 107–123.
- Matud MP (2004) Gender differences in stress and coping styles. *Personality and Individual Differences* 37(7): 1401–1415.
- Meinck S and Brese F (2019) Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-Scale Assessments in Education* 7(1): 1–23. DOI: 10.1186/s40536-019-0076-3.
- Mejía-Rodríguez AM, Luyten H and Meelissen MRM (2021) Gender differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. *International Journal of Science and Mathematics Education* 19(6): 1229–1250.

- Michelmore K and Sassler S (2016) Explaining the gender wage gap in STEM: Does field sex composition matter? *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2(4): 94.
- Möbius MM, Niederle M, Niehaus P, et al. (2022) Managing self-confidence: Theory and experimental evidence. *Management Science* 68(11): 7793–7817.
- Moore DA and Healy PJ (2008) The trouble with overconfidence. *Psychological Review* 115(2): 502–517.
- Moore DA and Klein WMP (2008) Use of absolute and comparative performance feedback in absolute and comparative judgments and decisions. *Organizational Behavior and Human Decision Processes* 107(1): 60–74.
- Mullis IVS, Martin MO, Foy P, et al. (2020) *TIMSS 2019 International Results in Mathematics and Science*. Chestnut Hill: Lynch School of Education and Human Development, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Musu-Gillette LE, Wigfield A, Harring JR, et al. (2015) Trajectories of change in Students' self-concepts of ability and values in math and college Major choice. *Educational Research and Evaluation* 21(4): 343–370.
- Nagy G, Trautwein U, Baumert J, et al. (2006) Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation* 12(4): 323–345.
- Németh E (1999) Gender differences in reaction to public achievement feedback. *Educational Studies* 25(3): 297–310.
- Neuschmidt O, Barth J and Hastedt D (2008) Trends in gender differences in mathematics and science (TIMSS 1995–2003). *Studies in Educational Evaluation* 34(2): 56–72.
- Niederle M and Vesterlund L (2007) Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics* 122(3): 1067–1101.
- Oakes J (1990) Opportunities, achievement, and choice: Women and minority students in science and mathematics. *Review of Research in Education* 16(1): 153–222.
- OECD (2019) *PISA 2018 Results (Volume II): Where All Students Can Succeed*. Paris: OECD Publishing.
- Reyes LH (1984) Affective variables and mathematics education. *The Elementary School Journal* 84(5): 558–581.
- Ridgeway CL (2011) *Framed by Gender: How Gender Inequality Persists in the Modern World*. New York, NY, US: Oxford University Press.
- Roberts TA and Nolen-Hoeksema S (1989) Sex differences in reactions to evaluative feedback. *Sex Roles* 21(11–12): 725–747.
- Robinson JP and Lubienski ST (2011) The development of gender achievement gaps in mathematics and Reading during elementary and middle school. *American Educational Research Journal* 48(2): 268–302.
- Ryan RM and Deci EL (2000) Self-Determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55(1): 68–78.
- Sax LJ, Kanny MA, Riggers-Piehl TA, et al. (2015) “But I’m not good at math”: The changing salience of mathematical self-concept in shaping women’s and men’s STEM aspirations. *Research in Higher Education* 56(8): 813–842.
- Schwardmann P and van der Weele J (2019) Deception and self-deception. *Nature Human Behaviour* 3(10): 1055–1061.
- Seymour E (1995) The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education* 79(4): 437–473.

- Shavelson RJ, Hubner JJ and Stanton GC (1976) Self-concept: Validation of construct interpretations. *Review of Educational Research* 46(3): 407–441.
- Smither JW, London M and Reilly RR (2005) Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology* 58(1): 33–66.
- Sterling AD, Thompson ME, Wang S, et al. (2020) The confidence gap predicts the gender pay gap among STEM graduates. *Proceedings of the National Academy of Sciences of the United States of America* 117(48): 30303–8.
- Sullivan A (2006) Students as rational decision-makers: The question of beliefs and attitudes. *London Review of Education* 4(3): 271–290.
- Suls J, Martin R and Wheeler L (2002) Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science* 11(5): 159–163.
- Swann WB, Rentfrow PJ and Guinn JS (2003) Self-verification: The search for coherence. In: Leary MR and Tangney JP (eds) *Handbook of Self and Identity*. New York, NY, US: The Guilford Press, 367–383.
- Tversky A and Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185(4157): 1124–1131.
- van de Werfhorst HG (2017) Gender segregation across fields of study in post-secondary education: Trends and social differentials. *European Sociological Review* 33(3): 449–464.
- Van Veldhuizen R (2022) Gender differences in tournament choices: Risk preferences, overconfidence, or competitiveness? *Journal of the European Economic Association* 20(4): 1595–1618.
- Varga J (ed) (2024) *A Közoktatás Indikátorrendszere 2023*. Budapest: HUN-REN KRTK, 1–338.
- Vinni-Laakso J, Guo J, Juuti K, et al. (2019) The relations of science task values, self-concept of ability, and stem aspirations among Finnish students from first to second grade. *Frontiers in Psychology* 10(JUL): 1–15.
- Wilkins JLM (2004) Mathematics and science self-concept: An international investigation. *Journal of Experimental Education* 72(4): 331–346.
- Wozniak D, Harbaugh WT and Mayr U (2014) The menstrual cycle and performance feedback Alter gender differences in competitive choices. *Journal of Labor Economics* 32(1): 161–198.

Author biography

Tamás Keller, PhD, is a senior researcher at the HUN-REN Centre for Social Sciences in Budapest, Hungary, and also affiliated with the Institute of Economics at the HUN-REN Centre for Economic and Regional Studies. His research focuses on education and social inequality.