



Gender and preferences at a young age: Evidence from Armenia[☆]



Karen Khachatryan^{a,*}, Anna Dreber^b, Emma von Essen^{c,d}, Eva Ranehill^e

^a Middlesex University Business School, The Burroughs, London NW4 4BT, United Kingdom

^b Department of Economics, Stockholm School of Economics, P.O. Box 6501, 11383 Stockholm, Sweden

^c Department of Economics and Business, Aarhus University, Fuglesangs allé 4, 8210 Aarhus, Denmark

^d TrygFonden's Centre for Child Research, Aarhus University, Fuglesangs allé 4, 8210 Aarhus, Denmark

^e Department of Economics, University of Zürich, Blümlisalpstrasse 10, 8006 Zürich, Switzerland

ARTICLE INFO

Article history:

Received 2 March 2014

Received in revised form

24 November 2014

Accepted 22 February 2015

Available online 20 March 2015

JEL classification:

C91

D03

J16

Keywords:

Competitiveness

Risk preferences

Altruism

Children

Gender differences

Experiment

ABSTRACT

We look at gender differences in competitiveness, risk preferences and altruism in a large sample of children and adolescents aged 7–16 in Armenia. Post-Soviet Armenia has few formal barriers to gender equality but is also characterized by a patrilineal kinship system and traditional gender roles. In contrast to research conducted in Western countries, we find that girls increase their performance more than boys in response to competition in a running task. We find no gender differences in the other three tasks we explore: skipping rope, a mathematical task, and a verbal task. We also find no difference in the willingness to compete in either the mathematical or the verbal task. In line with previous research, we find that boys are less altruistic and more risk taking than girls, and that the latter gap appears around the age of puberty.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Gender differences in economic preferences related to competitiveness, risk taking and pro-sociality are often brought forward as potential explanations to differences in economic choices and outcomes. While the literature shows that women are on average less competitive, less risk taking and more altruistic than men (see e.g., [Eckel and Grossman, 2008a,b](#); [Croson and Gneezy, 2009](#); [Bertrand, 2010](#); [Engel, 2011](#)), most of this research has been conducted on Western college students.

[☆] We thank Tore Ellingsen, Magnus Johannesson, Sten Nyberg, Jörgen Weibull, and participants in many conferences, workshops, and university seminars where this paper was presented for helpful comments and discussions. We also thank the Jan Wallander and Tom Hedelius Foundation (Svenska Handelsbankens Forskningsstiftelser), the Knut and Alice Wallenberg Foundation, and the Swiss National Science Foundation for financial support.

* Corresponding author. Tel.: +44 20 8411 4436; fax: +44 20 8411 5315.

E-mail addresses: k.khachatryan@mdx.ac.uk (K. Khachatryan), anna.dreber@hhs.se (A. Dreber), vonessen@econ.au.dk (E. von Essen), eva.ranehill@econ.uzh.ch (E. Ranehill).

A growing body of literature highlights the importance of studying different cultures, age groups and contexts in order to give us a more nuanced picture of gender differences in preferences, and their underlying mechanisms (e.g., Gneezy et al., 2009; Cardenas et al., 2012; Andersen et al., 2013). We report results from an experiment conducted among children and adolescents in Armenia, a developing country, on gender differences in preferences for competition, risk and altruism. Post-Soviet Armenia is today characterized by few formal barriers to gender equality. During the Soviet era female education and labor force participation was encouraged, mutual consent for marriage was implemented, and women's right to own land was established. However, the period since independence in 1991 has been described as a "slow return to patriarchy" (Kaser, 2008). Patriarchal norms and an unequal patrilineal and patrilineal kinship system favor boys over girls.¹ Despite the lack of formal barriers to equality, Armenia ranks 94th out of 136 countries in the Global Gender Gap Report of 2013 (Hausmann et al., 2013), with particularly low gender equality for health, survival and political empowerment. At the same time, the equality in educational attainment lies among the top 30 countries included in the report. For example, women comprise the majority of the population with higher education, but only 11% of the members of parliament and ministers are women.² In line with equal access to education, girls' academic achievement is high. According to international comparisons, girls have a higher mathematics achievement than boys (Mullis et al., 2012).³ Armenia is thus a country where girls' mathematics performance equals to or exceeds that of boys', but where traditional gender roles also seem to be strong. These contrasting aspects of Armenian society make Armenia an interesting country to study.

Our study comprises a sample of 824 children and adolescents aged 7–16 years in Armenia (in school grades 2–10). We focus on three aspects of behavior: competitiveness, risk preferences and altruism. These preferences are arguably important for economic outcomes, since they have been linked to important educational choices as well as outcomes in both labor markets and financial markets (see e.g., Zhang, 2013; Buser et al., 2014; Bonin et al., 2007). In addition, these are three areas in which gender differences are often observed. Moreover, these preferences have been explored in a variety of countries and settings.

We study competitiveness as (1) the performance response to competitive settings, as in Gneezy et al. (2003), in four tasks: running, skipping rope, a math task and a verbal task, and as (2) the willingness to compete, as in Niederle and Vesterlund (2007), in a math task and a verbal task. These tasks are associated with gender stereotypes, as confirmed by a survey executed after the experiment. This allows us to test if observed gender differences are context dependent and hence malleable. We measure risk preferences using the choice list method where subjects choose between different certain amounts and a gamble. Finally, we study altruism by having participants play a dictator game where the recipient is a charity (a well-known orphanage in Yerevan).

We find that boys and girls are equally competitive when looking at how performance changes in skipping rope, the math task and the verbal task, whereas girls are more competitive than boys in running. When it comes to the choice to compete, we find, contrary to most previous studies from Western countries, no gender gap in the willingness to compete in neither the math task nor the verbal task. Similarly, and in contrast to most previous literature, we find no gender differences in confidence, measured as relative performance beliefs, in the math and verbal tasks. Looking also at whether the gender gap varies across age groups, we follow Andersen et al. (2013) and focus on comparing prepuberty children with children around puberty by dividing the sample into "older" children (grade 7 or above) and "younger" children (grade 6 and below). We find no statistically significant difference in the gender gap across age groups in any of the competitiveness tasks included in the study. In the risk task, while the gender gap is significant in the sample of older children and not significant in the sample of younger children, it is significantly higher among older children than among younger children. This is mainly due to an increase of boys' risk taking with age, while girls' risk taking is constant across the age groups. We also find that girls are significantly more altruistic than boys. This applies to both older and younger children.

Compared to previous literature on children and adolescents, our most salient result is the lack of gender differences in competitiveness in the math task. When it comes to willingness to compete in mathematical tasks, most previous studies find that boys are more competitive than girls (Almas et al., 2012; Booth and Nolen, 2012a; Cardenas et al., 2012; Dreber et al., 2014; Sutter and Glatzle-Rutzler, 2015). For competitiveness in other types of tasks, there are many null results (Cardenas et al., 2012; Dreber et al., 2011; Samak, 2013), and also some evidence of boys being more competitive (Gneezy and Rustichini, 2004; Cardenas et al., 2012; Sutter and Glatzle-Rutzler, 2015).⁴

Our results lend further support to the view that context or culture may be an important determinant of gender differences in competitiveness. There are by now a number of studies that look at willingness to compete in a math task in different developed and developing countries. Cardenas et al. (2012), in line with previous research among adults, find that boys in Sweden are more willing to compete than girls, but find no gender difference in Colombia (classified as a developing

¹ It is worth noting that son-preference, once suppressed, is reviving in Armenia. Since the collapse of the Soviet Union, sex ratios at birth in Armenia have increased drastically and now are comparable to those found in current-day China and India (see e.g. Michael et al., 2013).

² For information on gender equality and educational achievement see country profiles in Hausmann et al. (2013), and for the percentage of women in parliament see Armenia at <http://www.ipu.org/parline-e/reports/2013-A.htm>, last accessed on October 19, 2014.

³ TIMSS 2011 International Results in Mathematics show a small and insignificant advantage of Armenian girls in comparison to Armenian boys in the fourth grade, and a larger, and significant, advantage in the eighth grade.

⁴ This is also true for the majority of previous research on adults. In most cases (Gneezy and Rustichini, 2004; Grosse and Reiner, 2010; Gunther et al., 2009; Kamas and Preston, 2012; Shurchkov, 2012) but not all (Wozniak et al., 2014) the gender gap in competitiveness among adults diminishes when the task performed is word related compared to, for example, solving mazes or simple mathematical tasks.

country). Similarly, [Almas et al. \(2012\)](#) find that boys are more competitive than girls in a Norwegian sample. Like Armenia (94), Colombia (35) is ranked below Sweden (4) and Norway (3) in gender equality according to the 2013 Global Gender Gap Report. In a similar vein, [Zhang \(2014\)](#) finds no gender gap in competitiveness among a group of Han Chinese (China ranks 69 in the 2013 Global Gender Gap Report). The above mentioned results, in combination with our null-result on willingness to compete in a math task in Armenia, suggest that it is important to compare different populations, such as developed and developing countries in order to better understand the gender gap in competitiveness. Competitiveness pertaining to the math domain also seems to be most important, in a sense that it has been found to correlate with many economic outcomes in contrast to the other domains of competitiveness explored here (e.g., [Niederle and Vesterlund, 2010](#)).

When it comes to risk preferences, our results are in line with most previous results, where boys are more risk taking. While [Harbaugh et al. \(2002\)](#) and [Almas et al. \(2012\)](#) find no gender gap, [Borghans et al. \(2009\)](#), [Cardenas et al. \(2012\)](#), [Dreber et al. \(2014\)](#), [Sutter et al. \(2013\)](#) find that boys are more risk taking than girls. There is some evidence that the gender gap in risk taking can also be influenced by culture or context (e.g., [Booth and Nolen, 2012b](#)). Results in this paper are in line with [Cardenas et al. \(2012\)](#) finding a larger gender gap in risk preferences in Colombia than in Sweden.

Regarding altruism, there is only a handful of studies on children and adolescents and the results are somewhat mixed. Some find that girls are more altruistic ([Dreber et al., 2014](#); [Gummerum et al., 2010](#)), some that there is no gender gap ([Almas et al., 2012](#); [Benenson et al., 2007](#); [Blake and Rand, 2010](#)), and one study finds that girls are less altruistic than boys ([Fehr et al., 2013](#)).

Further research is needed to shed light on the exact mechanisms contributing to the varying results across different countries. In Armenia, several policies during the Soviet era furthered gender-equality in an otherwise rather patriarchal society, and aspects such as female labor force participation may have had a long lasting impact on preferences related to competitiveness. This could explain why we find mainly null results for competitiveness together with one case where the “typical” gender gap is reversed. In a related vein, [Zhang \(2014\)](#) finds no gender gap in competitiveness among the Han Chinese, another population exposed to socialist policies that may have increased gender equality. However, it is hard to argue that Colombia has a similar history, yet [Cardenas et al. \(2012\)](#) find no gender differences in competitiveness there as opposed to in Sweden. It is also unclear why the “typical” gender gaps show up for risk preferences and altruism in this Armenian sample as well as in the Colombian sample in [Cardenas et al. \(2012\)](#). More work on this topic is thus needed: if enough countries are explored with similar protocols, it will eventually be possible to do a qualitative analysis of gender differences in preferences and their correlation with various gender equality variables at an aggregate level, such as educational attainment, labor force participation and political empowerment.

The rest of the paper is organized as follows. We present the experimental setup and procedures in Section 2 and the results in Section 3, and conclude with a discussion in Section 4.

2. Experimental design and procedures

The experimental design and procedures closely follow to that of [Cardenas et al. \(2012\)](#), as well as [Dreber et al. \(2014\)](#).⁵ The experiment took place in two nearby secondary schools in Yerevan, the capital of Armenia, during a four week period in April–May 2010. Overall 824 students aged 7–16 (in grades 2–10, 428 boys and 396 girls) participated in the study. The study consists of two parts: a physical education (PE) part and a classroom part.⁶

2.1. PE part: competition in running and skipping rope

The first part of the study was conducted during regular PE classes with students in grades 2–9 (aged 7–15).⁷ The students were simply told that they would participate in a series of physical tasks, without any more details. In this part we elicited competitiveness as the reaction to competition using two tasks: running and skipping rope. The students performed the tasks in a random order. Similar to a regular PE class, the students performed the tasks in the presence of their classmates. In this part, no compensation was awarded, as in [Gneezy and Rustichini \(2004\)](#) and [Dreber et al. \(2011\)](#).

Both tasks were performed in two rounds. In the first round, students performed the task individually. Hence, they were unaware of the existence of the second round when performing in the first round. After all students had completed all first-round tasks, they were matched with someone who had a similar performance to themselves, for each task. If more than two students obtained the same result, the matching was random. The students were made aware of the exact matching procedure, and were then asked to perform the task a second time in matched pairs. In addition, we let a separate, randomly selected, group of students perform the task individually in the second round, to serve as a control. This allows us to measure unobservable characteristics, such as fatigue in the second race, which may differ between genders.

⁵ In particular, for the younger children in grades 2–6 the experimental design and procedures are identical to that of [Cardenas et al. \(2012\)](#). The same design was implemented in order to facilitate potential future comparison with the data gathered in Colombia and Sweden.

⁶ To help run the experiment, we recruited and trained four female experimenters. For the first part, the experimenters worked together as a team overseeing different tasks, each specializing in one or two tasks so that, for example, running times were always recorded by the same experimenter. For the classroom part, the experimenters worked in teams of two. These experimenters were unaware of the gender focus of the study.

⁷ Note that we had access to a limited number of grade 10 students and they did not participate in the PE part of the study due to scheduling constraints. Detailed information on the number, age and gender of participants is available in the Supplementary Online Appendix.

The running task consisted of a 4 m × 13 m race (a total of 52 m). When competing in the race in the second round, paired students started at the same time and ran parallel to each other. The running task was the same for students of all grades. The skipping rope task, however, was different for younger and older students. Students in grades 2–6 skipped with a long rope that one teacher or experimenter turned. Here performance is measured by the number of jumps until the student first missed. Students in grades 7–9 skipped a regular rope for 1 min, with performance measured by the number of jumps in that one minute. When competing in skipping rope in the second round, two ropes were put next to each other and paired students were instructed to start skipping at the same time.

Our main measure of competitiveness in running and skipping rope is the absolute change in performance between the first and second rounds; this is the most common measure of the reaction to competition. However, we also report the results using relative performance change.

2.2. Classroom part: competitiveness in math and word search, risk preferences and altruism

The second part of the study took place during regular classes including students in grades 2–10 (aged 7–16). To begin with, the students were informed that they were taking part in an economic study and could earn prizes or money (depending on grade) by collecting *points* in all tasks. The students knew that more points corresponded to more prizes or more money which they could receive at the end of the study. After a discussion with the school principals it was decided that younger students – in grades 2–6 (aged 7–12) – were to be rewarded with pens, while older students – in grades 7–10 (aged 12–6) – were to be rewarded with money (see also Cardenas et al. (2012) vs. Dreber et al. (2014) for similar reward schemes).⁸

In this part we elicited competitiveness, risk taking, as well as altruistic preferences. Information about the procedure was here revealed sequentially and at no point in time did the students get feedback about their performance. The experiment was performed using pen and paper.

Inspired by Niederle and Vesterlund (2007), we elicited competitiveness as both the reaction to competition in terms of performance change as well as self-selection into competition using two tasks: a mathematical task and a verbal task. The math task consisted of simple arithmetic operations and varied in the level of difficulty depending on the grade. The verbal task was a word search puzzle, and the level of difficulty was the same for everyone. The order of these tasks was randomized for each class and grade, and detailed examples of these tasks can be found in Appendix A. Performance in both tasks was measured by the number of correct answers: the number of correctly solved exercises in the math task and the number of correct words found in the verbal task. Each of these tasks was performed three times, once in each of the first three stages. The students were sequentially informed of the incentive structure of each stage, which was as follows.

Stage 1: Piece rate pay. Students were asked to solve as many problems as possible in 2 min. They received 3 points for each correctly solved problem.

Stage 2: Tournament pay. Students were again asked to solve as many problems as possible in 2 min. They received 6 points for each correctly solved problem if they solved at least as many problems as a randomly selected student from their own class with whom they would be paired, otherwise they received 0 points.

Stage 3: Choice of tournament or piece rate pay. Students were asked to choose between the first stage piece rate and second stage tournament payment schemes and then solve as many problems as possible in 2 min.

Comparing performance under piece rate pay with performance under tournament pay gives as a measure of competitiveness as the reaction to competition, whereas the choice in the third stage gives as a measure of competitiveness as the willingness to compete. After each competitiveness task, we asked the students how many of their classmates they believed had solved more problems than they did. The same question was asked both for the first stage individual setting (piece rate) as well as the second stage competitive setting (tournament). The relative performance belief from the competitive setting is our measure of performance confidence.

After the competitiveness tasks, we elicited risk preferences. The risk task consisted of a list of six incentivized choices where students could choose between a safe option and a 50/50 gamble (executed as a coin flip in front of the students). The gamble was always the same and gave 10 points if a win and 0 points if a loss, while the safe option increased successively in points with each choice, ranging from 2 to 7.5. To resolve uncertainty and reward the students, six coins were flipped in front of the students at the end of the experiment, one for each choice. We employ two measures of risk preferences. The first relies on the unique point where participants switch from preferring the gamble to preferring the safe option. In this measure we exclude inconsistent students, i.e. those with multiple switching points. The second measure of risk preferences does not exclude these students and is defined by the number of times students choose the risky option over the safe one. We, thus, use the second as our main measure.

⁸ Throughout the results section we also analyze the sample as a whole even though we used two different reward schemes based on grade. We would like to thank a referee who pointed out that grouping participants with different motivations may not be appropriate because some of the results may be driven by the difference in motivation. We believe that even though the reward scheme could potentially impact the overall level of any behavior, it is unlikely that it affects gender differences which is the focus of our paper.

Table 1
Individual and competitive performance in running (seconds) and skipping rope (number of jumps)

Age group	Gender	Running task			Skipping rope task			N ^b
		Round 1 (Ind.)	Round 2 (Comp.)	SR test ^a (<i>p</i> -value)	Round 1 (Ind.)	Round 2 (Comp.)	SR test (<i>p</i> -value)	
Grades 2–6	Boys	17.39	16.98	<0.001	5.92	6.39	0.340	154/153
	Girls	18.95	18.32	<0.001	8.38	7.12	0.692	147/145
Grades 7–9	Boys	15.13	14.92	<0.001	67.10	72.27	<0.001	100/96
	Girls	16.94	16.27	<0.001	70.92	79.45	<0.001	91/88
All grades	Boys	16.50	16.17	<0.001	29.66	31.79	0.002	254/249
	Girls	18.19	17.54	<0.001	32.50	34.44	0.016	238/233

^a Wilcoxon matched pairs signed-rank test, testing whether (Ind.)=(Comp.), henceforth SR test in the tables.

^b 5 boys and 5 girls took part in the running task, but did not complete the skipping rope task.

After the risk task, we elicited altruistic behavior through a dictator game. The students were asked to allocate 100 points between themselves and a well-known Armenian charity organization (the Zeitoun Orphanage in Yerevan). Students were informed that points allocated to the charity would be converted to gifts and money, and sent by the experimenters to the orphanage. The amount donated to the charity is our measure of altruism.

Finally, the students were asked to fill out a survey including demographic questions and questions on gender stereotypes associated with the different tasks. At the end of the classroom part, points were counted and converted to prizes and money and the students were given their respective rewards.

In sum, the experimental design allows us to analyze (1a) competitiveness as performance change in running, skipping rope, a mathematical task and a verbal task, (1b) competitiveness as willingness to compete in a mathematical task and a verbal task, (2) risk preferences through incentivized choices, and (3) altruism via a dictator game with a charity recipient.⁹

3. Results

In this section we first look at the gender gap in competitiveness, through both the performance change as a response to competition in all four tasks and the willingness to compete in the two classroom tasks. For the two physical tasks, where the gender of the counterpart was known, we also study whether the gender composition of the competing pairs matters for the individual response to competition. We then look at gender differences in risk taking, and finish with altruism. Throughout this section, in addition to reporting results based on the whole sample, for each experimental task we also explore potential gender and age interactions. We here follow Andersen et al. (2013) and focus on comparing prepuberty children with children around puberty, by dividing the sample into “older” children (grade 7 and above) and “younger” children (grade 6 and below).¹⁰ To examine whether the gender gap varies over time, we use difference-in-difference regressions, summarily reported in Table 7 in Appendix C. In addition, we present a brief robustness analysis. The analysis based on the answers from the survey conducted in the classroom is presented in Appendix B.

Throughout the paper all tests of (gender) differences in means are analyzed using the non-parametric Wilcoxon–Mann–Whitney and χ^2 proportions tests.¹¹

3.1. Competitiveness

3.1.1. Competitiveness as a reaction to competition in running and skipping rope

Table 1 shows average performance in the two physical tasks under individual and competitive incentives across age groups, for boys and girls separately. We observe that the gender gap in individual performance largely corresponds to the gender stereotype of the task. Whereas boys run significantly faster ($p < 0.001$ for all pairwise comparisons), girls perform significantly better in the skipping rope task (p -values for a gender difference are 0.023, 0.087, and 0.007 for younger, older and all students, respectively).¹² In the second round, when participants compete against each other, both boys and girls of all ages increase their performance significantly in running. In the skipping rope task, however, only the older children are competitive in terms of increasing their performance in the second round.

⁹ The students also participated in a cooperation task in the first part of the study (PE part). In the second part of the study (classroom part), the students also participated in a public goods game and a time preference task. These measures are beyond the scope of the current paper.

¹⁰ This largely corresponds to having a cutoff at the age of 13 as in Andersen et al. (2013), since the median age in grade 7 at the time of our experiment was 13 years. Also, with this division, participants in each age group were provided with the same incentives.

¹¹ We have also compared whether the distributions for each reported variable differ between boys and girls using a Kolmogorov–Smirnov test. The results are similar to those reported for mean values, unless otherwise noted.

¹² p -values come from the pooled sample of competition and control groups, since all conditions were identical in the first round.

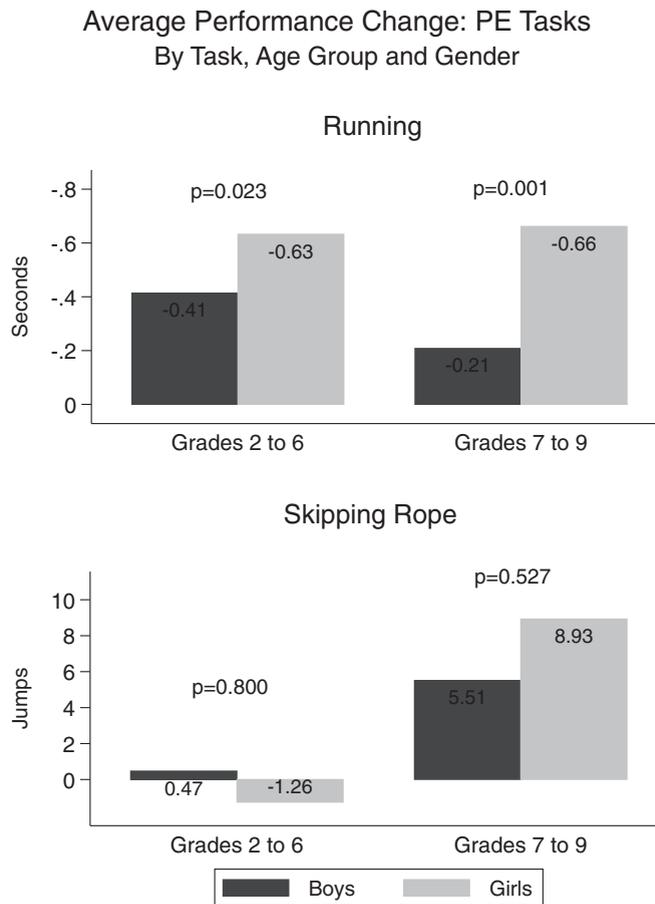


Fig. 1. Gender and competitiveness in running and skipping rope.

Comparing the absolute change in performance between the first and the second rounds, we find that even though both boys and girls of all ages run faster in the second round, girls increase their performance significantly more than boys, see Fig. 1 (for the whole sample: $p < 0.001$). However, in the skipping rope task we do not find any gender differences in competitiveness (for the whole sample: $p = 0.917$).¹³

Comparing the size of the gender gaps in performance change across age groups, we do not find any significant differences in neither running ($p = 0.341$) nor skipping rope ($p = 0.107$).

In both running and skipping rope tasks, the students could observe each other while performing in the second round. In Table 2 we therefore explore whether the response to competition varies with the gender of the opponent in our sample. In running, both boys and girls improve their performance significantly more when competing against a girl versus a boy (boys: $p = 0.002$; girls: $p = 0.006$), and this difference is mainly driven by the older students.¹⁴ However, we find no significant opponent gender effects in the skipping rope task.

Summarizing, we have the following result:

Result 1. *Girls are significantly more competitive than boys in the running task, while there is no significant gender difference in the skipping rope task.*

¹³ We have also conducted the same analysis using two measures of relative performance. We calculate the performance change for each individual relative to the class mean, since the classmates can see each other when they perform in this part of the experiment. Further, older and younger children conducted different types of the skipping rope tasks. This measure of relative performance change takes this into account. The second measure is the relative change in individual performance over the two rounds, i.e. $(\text{performance in round 2} - \text{performance in round 1}) / (\text{performance in round 1})$. For both relative measures used, the main results are qualitatively similar (for younger children in the running task while the first relative measure of performance change is significant at conventional levels, the second relative measure is significant at 10% in a two sided test, and at 5% in a one sided test). Furthermore we have also checked if there are any differences in the variance of performance change between boys and girls. While for the running task there are no significant differences for any age group ($p < 0.05$ for all pairwise comparisons, Levene's test), in the skipping rope task girls have a significantly higher variance: at the median $p = 0.033$ for younger, $p = 0.086$ for older and $p = 0.016$ overall, Levene's test. However, this gender difference in variance also vanishes when we look at the relative performance change (all $p > 0.5$).

¹⁴ In the sample of younger children, while this difference is not significant for boys ($p = 0.378$), it is marginally significant for girls ($p = 0.073$).

Table 2
Impact of opponent's gender on competitiveness

Sample of	Running against ^a				Skipping rope against ^b			
	Boys	Girls	p-value	N	Boys	Girls	p-value	N
Boys	-0.25	-0.49	0.008	162/91	2.12	2.74	0.558	129/118
Girls	-0.46	-0.76	0.006	94/144	2.57	2.60	0.996	118/118

^a Average performance change (seconds).

^b Average performance change (number of jumps).

Table 3
Individual and competitive performance in the mathematical and verbal tasks

Age group	Gender	Math task ^a			Verbal task ^b			N
		Stage 1 (Ind.)	Stage 2 (Comp.)	SR test ^a (p-value)	Stage 1 (Ind.)	Stage 2 (Comp.)	SR test (p-value)	
Grades 2–6	Boys	17.44	20.83	<0.001	5.31	5.56	0.071	224 ^c
	Girls	16.36	20.12	<0.001	5.80	6.09	0.137	204
Grades 7–10	Boys	8.12	10.32	<0.001	9.20	9.74	0.112	164
	Girls	7.96	8.29	<0.001	10.61	11.61	0.003	173
All grades	Boys	13.50	16.18	<0.001	6.95	7.33	0.018	388
	Girls	12.51	15.62	<0.001	8.01	8.62	0.001	377

^a Average number of correctly solved math exercises.

^b Average number of correct words found in the word search puzzle.

^c One boy did not participate in the second stage of the verbal task.

3.1.2. Competitiveness as a reaction to competition in math and word search

Table 3 shows average performance in the two classroom tasks under individual (piece rate) and competitive (tournament) incentives across ages, for boys and girls separately. We observe, again, that the gender gap in the first stage piece rate performance largely corresponds to the gender stereotype of the task. Girls perform significantly better in the verbal task ($p < 0.05$ for all pair-wise comparisons). However, while boys solve slightly more exercises in the math task, this difference is not significant ($p > 0.10$ for all pairwise tests). In the second stage, when students compete against each other, both boys and girls of all ages are competitive and react to competition by increasing their performance significantly in both tasks.¹⁵

In Fig. 2 we compare the change in performance between the first and the second stages in the two classroom tasks. We find that boys and girls of all ages react similarly to competition and improve equally in both tasks (p -value for the whole sample, comparing boys vs girls is 0.195 for math and 0.328 for the verbal task).¹⁶ Moreover, there is no difference in the size of the gender gap across age in both tasks (math: $p = 0.771$; word: $p = 0.428$).

Summarizing, we have the following result:

Result 2: Boys and girls do not differ significantly in terms of performance change when reacting to competition in both a mathematical and a verbal task.

3.1.3. Competitiveness as willingness to compete in math and word search

Contrary to the two physical tasks, both the math task and the verbal task included a third stage, where participants could choose to perform the task under either piece rate incentives or competitive incentives. In the math task, 54% of the boys and 52% of the girls choose to compete, and in the verbal task the share is 57% for boys and 56% for girls. Neither gender difference is significant (math task: $p = 0.708$, verbal task: $p = 0.697$; χ^2 test for equality of proportions).

In Fig. 3 we explore competitive choices of boys and girls in each task for the two age groups separately. While we do not find any gender difference within any age group, 61% of younger and 43% of older students choose to compete in math, and this difference is significant. This is also true for each gender over age ($p = 0.001$ for all pairwise comparisons, χ^2 test). In the verbal task, 55% of younger and 59% of older students choose to compete. However we do not find any significant gender or age differences in this task ($p > 0.19$ for all pairwise comparisons, χ^2 test).

Summarizing, we have the following result:

Result 3: Boys and girls do not differ significantly in their willingness to compete.

Comparing this result to those of Cardenas et al. (2012) who use similar tasks, it appears that Armenian children are more competitive than both Colombian and Swedish children, in particular if we compare Armenian girls with Swedish girls. Among Colombian children, 35% of the boys and 32% of the girls choose to compete in the math task and 26% of the

¹⁵ In the second stage, again, boys perform better in the math task and girls perform better in the verbal task, but only in the verbal task the gender difference is significant (math: $p = 0.384$, verbal: $p = 0.001$).

¹⁶ This finding is also robust to using relative measures of performance change. Furthermore, we have also checked whether there are gender differences in the variance of performance change for both tasks. While girls have a higher variance in performance change in the verbal task (all $p < 0.01$, Levene's test), this difference vanishes when we look at relative performance change (all $p > 0.19$, Levene's test).

Average Performance Change: Classroom Tasks By Task, Age Group and Gender

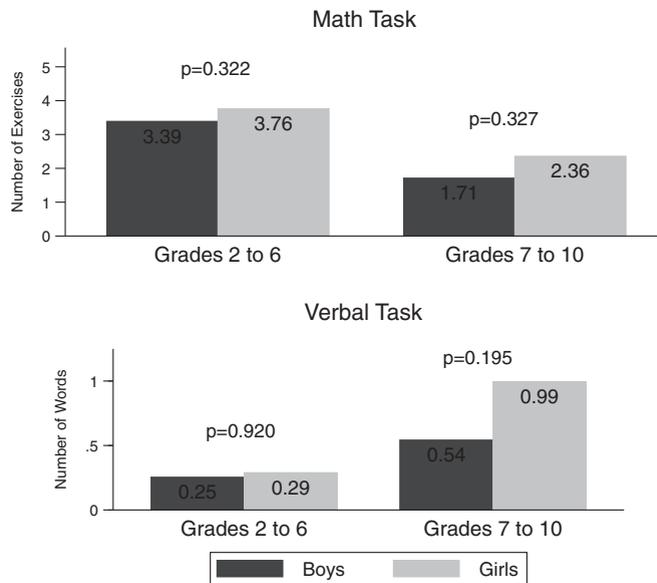


Fig. 2. Gender and competitiveness in math and word search.

boys and 29% choose to compete in the verbal task. Among Swedish children, the corresponding numbers are 44% of the boys and 19% of the girls in the math task and 39% of the boys and 27% of the girls in the verbal task.

3.1.4. Relative performance beliefs in math and word search

As a measure of confidence related to performance in the classroom tasks we asked the children and adolescents about relative performance beliefs, both in the first stage piece rate setting and in the second stage tournament setting. In contrast to some previous research (e.g., Niederle and Vesterlund, 2007) – and interesting in relation to our null results on gender

Share of Individuals Willing to Compete By Task, Age Group and Gender

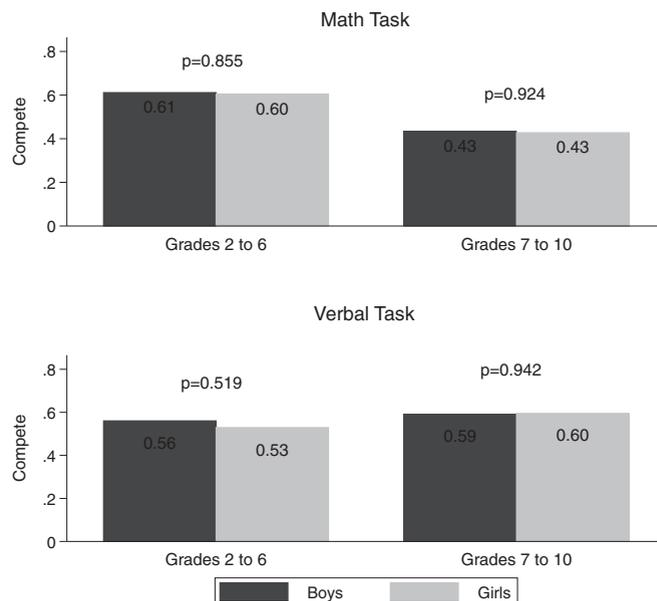


Fig. 3. Gender and the choice to compete.

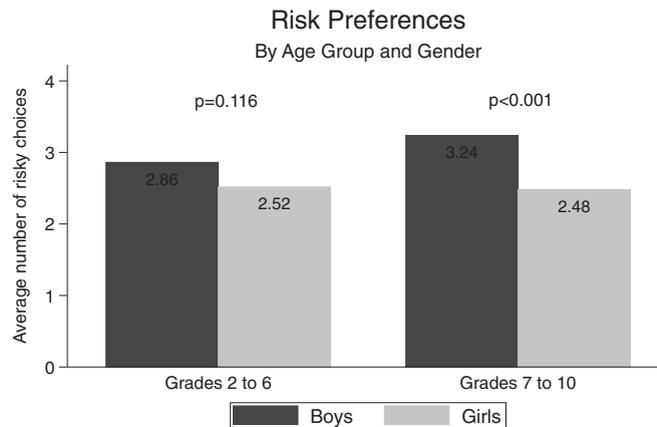


Fig. 4. Gender differences in risk preferences.

differences in competitiveness in the classroom tasks – we find no gender differences in relative performance beliefs for either setting (piece rate math: $p = 0.286$; piece rate verbal: $p = 0.163$; tournament math: $p = 0.594$; tournament verbal: $p = 0.633$). Moreover, our main results pertaining to gender and competitiveness (both performance change and willingness to compete) do not change if we control for relative performance beliefs in a parametric regression framework.¹⁷ This result is in line with that of Cardenas et al. (2012), who also did not find that performance beliefs mattered for willingness to compete in Colombia.

3.2. Risk preferences

In this section we test whether there are gender differences in risk preferences in our sample. We measure risk preferences from six incentivized choices in the classroom. In our sample of 762 students that took part in the risk task, about 19% are inconsistent in their choices, i.e. have multiple switching points. (There is no gender difference in being inconsistent in the whole sample, or any of the age groups considered: all $p > 0.25$, χ^2 equality of proportions test, see Supplementary Online Appendix.) Inconsistent participants are excluded if we look only at the unique switching point. We therefore use the number of risky options chosen as our main measure of risk preferences. We find that, on average, boys choose the risky option significantly more often than girls (boys: 3.02, girls: 2.50, $p < 0.001$). On average boys make 21% more risky choices compared to girls.

In Fig. 4 we display gender differences in risk preferences in the two age groups separately. While boys make significantly more risky choices than girls do, this gender difference does not reach statistical significance for younger children. It appears that in our sample a gender difference in risk taking arises around the age of puberty, and this is mainly due to increased risk taking of boys with age, see also Table 7 in Appendix C. Using the other risk measure based on the unique switching point, and thus excluding inconsistent participants, our results are qualitatively the same. Testing whether the gender gap in risk taking is larger among older children than younger ones, we find that while this difference is marginally significant with our main measure of risk ($p = 0.06$), it is significant for the other measure that excludes inconsistent participants ($p = 0.04$). However, we should also note that the distribution of risk preferences is remarkably different across our age groups, see Supplementary Online Appendix.

Summarizing, we have the following result:

Result 4: Boys are more risk taking than girls, and this gender gap gets bigger in adolescence.

These results on risk preferences are in line with those of Cardenas et al. (2012), who find that Colombian boys on average take 40% more risk than girls, with the corresponding number in Sweden being 15%.

3.3. Altruism

In this section we look at gender differences in altruism as measured via donations in a dictator game. We find that girls give significantly more than boys ($p < 0.001$), and this is the case for both age groups, see Fig. 5. Girls donate on average 60.16 points whereas boys donate 51.48 points out of 100 to the charity organization (an orphanage). The modal allocation is the 50–50 split for the sample as a whole, as well as for both genders and age groups.

¹⁷ For both tasks, when regressing willingness to compete on female, age, actual performance in the tournament and confidence as measured via performance beliefs from the tournament, performance is significant and positive while confidence is significant and negative. There are also no significant gender and confidence interactions, see Supplementary Online Appendix.

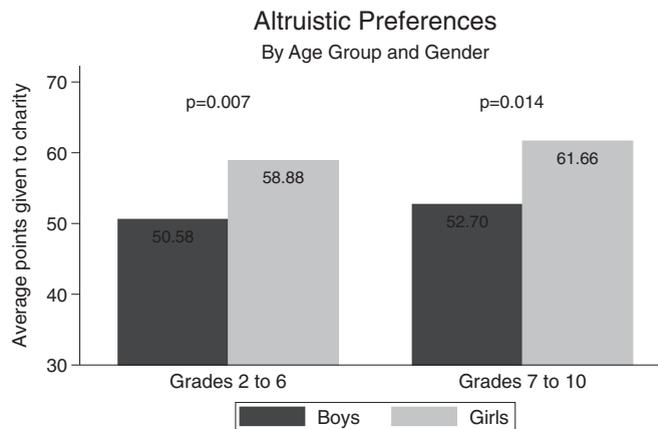


Fig. 5. Gender differences in altruism.

Testing whether the gender gap in altruism is bigger among the older students, we do not find a significant difference ($p=0.89$). Hence, we have the following result:

Result 5: *Girls are significantly more altruistic than boys.*

3.4. The relationship between competitiveness, risk preferences, and altruism

We also explore the relationship between competitiveness, risk taking, and altruism, since the three areas are often related, yet are three separate concepts, and gender differences are often found in all three preferences. A correlation analysis between all the behaviors we examine shows that while competitiveness and risk taking are related, our measure of altruism is neither related to competitiveness nor to risk taking, see Supplementary Online Appendix.¹⁸

Like many others, e.g. Niederle and Vesterlund (2007), Niederle and Yestrumskas (2008), we find a positive relationship between risk taking and willingness to compete in math and word search, indicating that individuals who are willing to compete also tend to be more risk taking, and vice versa ($p < 0.001$ for both tasks).¹⁹ Studying the sample split by gender we find a similar relationship.

However, when we look at the relationship between risk taking and willingness to compete for younger and older students separately, we find a different relationship for older students. While both older boys and older girls display a positive and significant relationship between risk taking and willingness to compete in the verbal task (boys: $p=0.029$, girls: $p=0.007$, overall: $p < 0.001$), we do not find any relationship between risk taking and willingness to compete in the math task (boys: $p=0.489$, girls: $p=0.204$, overall: $p=0.295$). This shows the importance of studying different age groups in order to increase our understanding of the relationship between competitiveness and risk taking.

3.5. Robustness analysis

3.5.1. Control group

It is possible that boys and girls have different physical capacity, learning curves or differ in other ways that influence their sequential performance differently. If, for example, girls get tired faster than boys, such that it negatively influences their second round performance, this would falsely be interpreted as a lower response to competition among girls in our design. We therefore have a separate group of students act as a control group by performing the task individually in the second round of the two physical tasks. The control group includes 135 students in the running task (69 boys and 66 girls) and 132 students in the skipping rope task (67 boys and 65 girls). While students on average do not improve their performance in the second round, the difference in change of performance between boys and girls is not significant in either of the two tasks, see Table 5 in Appendix C.

3.5.2. Age and the gender gap in preferences

Throughout the results section we have grouped children into age categories with a cutoff at grade 7. We have also tested whether our main results are robust to choosing a cutoff at grades 6 and 8. Using different cutoff points does not

¹⁸ We have also tested whether the distributions of donations differ between those who choose to compete and those who do not, across tasks, age and gender. For both tasks none of these (pairwise) comparisons is significant, all $p > 0.20$.

¹⁹ The p -values come from testing equality of distributions of risk preferences between those who chose competition and those who did not, using a Kolmogorov–Smirnov test. This is the case for both risk measures.

change our results for competitiveness.²⁰ However, if we use grade 8 as a cutoff point (and thus have disproportionately more participants below that threshold), the gender gap in altruism is significantly larger among older participants, while the difference in gender gaps in risk preferences vanishes with our main measure of risk taking and becomes marginally significant only with our secondary measure of risk preferences that excludes inconsistent participants ($p=0.08$). Using grade 6 as a cutoff point renders these differences insignificant as well.

3.5.3. Age as a continuous variable

So far we have examined age effects through age group dummies. Since we have a large range of age for students, for the four competitiveness tasks we have also conducted parametric regression analyses to test the robustness of our findings and to see whether the gender effect increases or decreases with age, using age as a continuous variable instead of age group dummies, along with other control variables.²¹ When it comes to gender differences in competitiveness our main results do not alter in a qualitative way. All regression tables are available in a Supplementary Online Appendix.²²

3.5.4. Age variation within a school grade

In our data set we see some variation in age within each grade. If the age at school start, or dropouts and repetition of a grade, is correlated with gender, this may bias our results if, for example, being slightly older also means having a physical and academic advantage in the competition. In our sample girls are slightly older compared to boys. The average age difference is small, about a month, but statistically significant. We thus conduct a robustness check of our main results using parametric OLS regressions including a variable capturing the distance between each student's age and the mean age within the respective class the student attends (see Table 6 in Appendix C). We further cluster on class to control for class specific settings. Including this variable does not change our results on gender differences in competitiveness, risk preferences or altruism in a qualitative way.

4. Discussion and conclusion

There are relatively few studies on gender differences in developing countries. In this paper we study whether there are gender differences in preferences related to competitiveness, risk taking and altruism among a large sample of children and adolescents in Armenia, a transitioning country with high gender inequality and distorted sex-ratios. We find mixed results when it comes to gender differences in preferences. Unlike most other studies in Western countries, the girls in our sample are at least as competitive as the boys. They are equally competitive in three out of four tasks and more competitive than boys in a fourth task. We also find that girls are on average less risk taking and more altruistic than boys, results in line with most other studies.

Armenia differs from many of the countries previously explored in the literature on gender differences in economic preferences. In contrast to these countries, Armenia is a developing country, and in important respects one of the least gender equal countries explored so far. Only India (rank 114), where Andersen et al. (2013) study two different matrilineal and patrilineal societies, scores lower than Armenia (rank 94) according to the Global Gender Gap Index of 2013 (Hausmann et al., 2013). Although Armenia has no legal barriers to gender equality, traditional gender roles are emphasized and the country has a patrilineal kinship system. To what extent similar results would be found in other former member states of the Soviet Union is a natural extension of this research.

We find the lack of a gender difference in willingness to compete in the math task to be the most striking result from this study. As far as we know, all studies exploring willingness to compete in mathematical tasks using Western samples find that men and boys compete more. The results in developing countries are mixed. Cardenas et al. (2012) find no gender gap in Colombia. Zhang (2014) finds similar null-results among the Han Chinese. At the same time, however, Zhang (2014) finds that boys are significantly more competitive in two other more traditional groups in China, the Yi and the Mosuo. To some extent, gender differences in competitiveness in math tasks may thus be more prevalent in Western societies. More studies are needed to explore the mechanisms behind why gender differences in competitiveness show up in some societies but not others.

Our sample further spans a large age range for children, enabling us to explore whether gender differences are more or less pronounced in different age groups. We find an indication that the gender gap in risk preferences is larger among the

²⁰ Students aged 7–12 below grade 7 were rewarded differently than the older students. As a referee pointed out, this could potentially induce differences in results when comparing older and younger children, especially when participants with different reward schemes are grouped together, that have nothing to do with age per se. However, we believe that it is unlikely that this would affect boys and girls differently, and thereby induce a change in the size of the gender gap across ages, which is the main focus of this analysis. We have also used ages 12, 13 and 14 as cutoff points, and the results for competitiveness are qualitatively the same.

²¹ Note that we have collected age data (year and month of birth) through the survey at the end of the classroom part of the experiment. About 5% of the students did not (fully) answer this question. We also lack age data for those students who took part in the PE part of the experiment but were absent on the day the classroom part took place. In subsequent analysis, wherever we use exact age data we ignore these individuals. Imputing the age from school grade for these students and using that variable instead does not qualitatively change our results.

²² For the classroom tasks, we have also repeated the same analysis for younger and older students separately. The only difference we find is in the willingness to compete regressions, where it is confirmed that risk taking does not explain competitive choices in the math task for older students.

older children (aged 13 and above). This increase is mainly due to an increased risk appetite among older boys, who appear to become more risk taking around the age of puberty, perhaps suggesting that the gender gap in risk taking is to an extent related to hormonal changes during this period (Apicella et al., 2008, 2014; Sapienza et al., 2009; for a conflicting view, see Zethraeus et al., 2009).

In sum, our results provide further evidence that it is important to examine the role of cultures and contexts in order to increase our understanding of gender gaps in economic preferences, both in developed and developing countries.

Appendix A.

In this appendix we present examples of the competitiveness tasks in the classroom part of the study: the mathematical and verbal tasks.

Table 4 shows a few examples of the mathematical exercises that the children solved for various grades. Students in grades 2–5 had only to sum a random sequence of two 1- and 2-digit numbers, while students in grades 6 and 7 had to both add and subtract a random sequence of two 2-digit numbers. Students in grades 8–10 had to both add and subtract a random sequence of three 2-digit numbers. All the numbers and the mathematical operations were randomly generated to insure that the level of difficulty of the mathematical task was the same throughout all the stages of the experiment for each of the grade categories.

Fig. 6 shows an example of the verbal task that was used in the experiment – a word search puzzle: these are letters in the Armenian alphabet. The students had to find and circle words in any direction on a straight line, including diagonally. The puzzles contained many words of different lengths and were the same for students of all grades. Since these word search

Table 4
Examples of the mathematical tasks for various grades

Grades 2 and 3	Grades 4 and 5	Grades 6 and 7	Grades 8–10
1 + 12 = ...	82 + 18 = ...	93 + 67 = ...	96 + 93 + 3 = ...
3 + 5 = ...	48 + 10 = ...	63 – 38 = ...	33 – 9 – 85 = ...
11 + 4 = ...	47 + 14 = ...	2 – 38 = ...	83 + 97 + 14 = ...
17 + 18 = ...	39 + 6 = ...	71 + 52 = ...	31 – 39 + 28 = ...
13 + 8 = ...	65 + 7 = ...	58 – 72 = ...	47 – 11 + 5 = ...
9 + 14 = ...	99 + 1 = ...	51 + 27 = ...	63 + 17 – 72 = ...
10 + 23 = ...	68 + 16 = ...	89 – 46 = ...	9 – 41 – 75 = ...

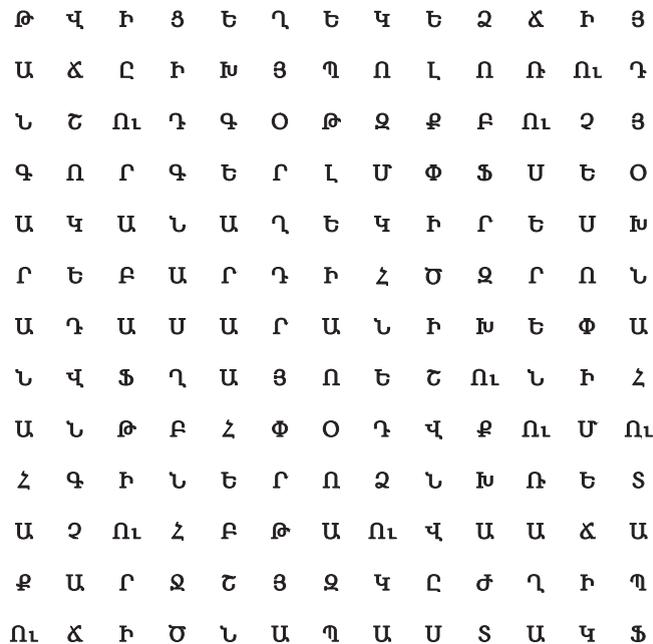


Fig. 6. An example of the verbal task: word search puzzle.

puzzles were not randomly generated by a computer there might have been slight differences in the difficulty of the verbal task in different stages of the experiment. We used the same puzzle within each stage of the experiment for all the students.

Appendix B. Additional analysis

At the end of the experiment we administered a survey, in which, among other things, we elicited perceptions of how boyish/girlish the students considered running, skipping rope, the math task and the verbal task to be. We further asked how boyish/girlish they considered competing in these tasks to be. We used an eleven point scale from 0 to 10 where a lower number indicates rating the task as more girlish and a higher number as more boyish (0 corresponded to being very girlish, 5 corresponded to being gender neutral, 10 corresponded to being very boyish). We used a similar scale to elicit how important the students considered competing against a boy and against a girl to be in two different questions (with 0 being not at all important and 10 being very important).

B.1. Are tasks perceived as gendered?

Boys perceive running to be significantly more boyish than girls do (boys: 8.78, girls: 6.69; $p < 0.001$). Boys rate skipping rope as more gender neutral, while girls consider it to be more girlish and the difference in ratings is statistically significant (boys: 6.05, girls: 3.85; $p < 0.001$). When it comes to the math task, boys consider it to be more boyish, while girls consider it to be more gender neutral (boys: 7.48, girls: 5.36; $p < 0.001$). When it comes to the verbal task (word search), both boys and girls perceive it to be closer to gender neutral but still differ in their opinions (boys: 6.64, girls: 4.30; $p < 0.001$).

B.2. Is competing important?

Boys rate competing as more fun compared to girls (boys: 8.80; girls: 8.14; $p < 0.001$), but boys and girls rate winning a competition as equally important (boys: 8.40; girls: 8.21; $p = 0.231$). Boys believe it is more important to win against a boy than against a girl (against a boy: 8.90, against a girl: 5.95; $p < 0.001$), and there is some evidence that girls share this view though the difference is only marginally significant (against a boy: 7.78, against a girl: 7.51; $p = 0.079$). Interestingly, this does not correspond to what we observe in actual performance change in running and skipping rope, where the gender of the opponent is known. In both tasks, both boys and girls improve their performance more when competing against a girl, see the main text.

Appendix C. Robustness analysis – tables

Tables 5–7.

Table 5

Control group: running and skipping rope performed individually twice.^a

Age group	Gender	Running task			Skipping rope task			N
		Round 1 (Ind.)	Round 2 (Comp.)	SR test (p-value)	Round 1 (Ind.)	Round 2 (Comp.)	SR test (p-value)	
Grades 2–6	Boys	17.69	17.93	0.034	5.93	7.82	0.051	45/45
	Girls	18.65	18.61	0.456	13.94	11.41	0.536	35/34
Grades 7–9	Boys	14.62	14.43	0.424	55.50	64.45	0.023	24/22
	Girls	16.15	15.93	0.667	71.90	73.13	0.281	31/31
All grades	Boys	16.62	16.71	0.269	22.21	26.42	0.004	69/67
	Girls	17.45	17.35	0.806	41.58	40.85	0.486	66/65
	Total	17.32	16.83	0.325	31.75	33.53	0.011	135/132

^a There are no significant gender differences in performance change in either task and age group. When testing for a gender difference in performance change $p > 0.10$ for all pairwise comparisons.

Table 6

Age difference: OLS regressions, dependent variables are performance change in competitiveness tasks and willingness to compete

Independent variables	(1) Running	(2) Skipping	(3) Math	(4) Word	(5) Math choice	(6) Word choice
Female	−0.257*** (0.0629)	1.670 (1.990)	0.480 (0.525)	0.225 (0.254)	−0.0210 (0.0335)	−0.0175 (0.0358)
Age difference	−0.165 (0.109)	2.378 (2.201)	−0.643 (0.390)	0.0960 (0.324)	0.0344 (0.0431)	0.0007 (0.0418)
Observations	420	412	727	727	727	726
R-squared	0.032	0.007	0.005	0.001	0.001	0.000

Robust standard errors in parentheses, clustered on class.

* p<0.1.

** p<0.05.

*** p<0.01.

Table 7

Age and the gender gap in competitiveness, risk and altruism

Independent variables	(1) Running	(2) Skipping	(3) Math	(4) Word	(5) Math choice ^a	(6) Word choice	(7) Risk 1	(8) Risk 2	(9) Altruism
Age Group ^b	0.206 (0.178)	5.040 (2.956)	−1.675 (1.027)	0.288 (0.490)	−0.174*** (0.0573)	0.0482 (0.0838)	0.377 (0.265)	0.549 (0.418)	2.118 (5.120)
Female	−0.219* (0.107)	−1.733 (1.481)	0.371 (0.544)	0.0348 (0.244)	−0.007 (0.0312)	−0.0157 (0.0340)	−0.344** (0.140)	−0.449** (0.211)	8.294** (3.057)
Female × Age group	−0.235 (0.241)	5.154 (3.069)	0.274 (0.934)	0.417 (0.519)			−0.417* (0.214)	−0.736** (0.338)	0.664 (4.770)
Constant	−0.414*** (0.111)	0.471 (0.898)	3.388*** (0.852)	0.254 (0.169)			2.861*** (0.225)	4.507*** (0.354)	50.58*** (3.616)
Observations	491	482	765	765	765	764	762	620	763
R-squared	0.039	0.058	0.024	0.008			0.024	0.030	0.018

Robust standard errors, clustered on class, in parentheses

* p<0.1.

** p<0.05.

*** p<0.01.

^a The math choice and word choice columns report the total marginal effect of age and female from probit regressions. The marginal effect of the interaction term only is not significant.^b Age group is a dummy variable taking on value 1 if grade 7 or above, and 0 otherwise.

Appendix D. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jebo.2015.02.021>.

References

- Almas, I., Cappelen, A.W., Salvanes, K.G., Sorensen, E.O., Tungodden, B., 2012. Explaining Gender Differences in Competitiveness: Family Matters (unpublished).
- Andersen, S., Ertac, S., Gneezy, U., List, J.A., Maximiano, S., 2013. Gender, competitiveness and socialization at a young age: evidence from a matrilineal and a patriarchal society. *Rev. Econ. Stat.* 95, 1438–1443.
- Apicella, C.L., Dreber, A., Campbell, B.C., Gray, P., Hoffman, M., Little, A.C., 2008. Testosterone and financial risk preferences. *Evol. Hum. Behav.* 29, 384–390.
- Apicella, C.L., Dreber, A., Mollerstrom, J., 2014. Salivary testosterone change following monetary wins and losses predicts future financial risk-taking. *Psychoneuroendocrinology* 39, 58–64.
- Benenson, J.F., Pascoe, J., Radmore, N., 2007. Children's altruistic behavior in the dictator game. *Evol. Hum. Behav.* 28, 168–175.
- Bertrand, M., 2010. New perspectives on gender. *Handbook of Labour Economics*, vol. 4b. Elsevier, pp. 1545–1592 (Chapter 17).
- Blake, P.R., Rand, D.G., 2010. Currency value moderates equity preference among young children. *Evol. Hum. Behav.* 31, 210–218.
- Bonin, H., Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2007. Cross-sectional earnings risk and occupational sorting: the role of risk attitudes. *Labour Econ.* 14, 926–937.
- Booth, A.L., Nolen, P.J., 2012a. Choosing to compete: how different are girls and boys? *J. Econ. Behav. Org.* 82, 542–555.
- Booth, A.L., Nolen, P.J., 2012b. Gender differences in risk behaviour: does nurture matter? *Econ. J.* 122, F56–F78.
- Borghans, L., Golsteyn, B., Heckman, J.J., Meijers, H., 2009. Gender differences in risk aversion and ambiguity aversion. *J. Eur. Econ. Assoc.* 7, 649–658.
- Buser, T., Niederle, M., Oosterbeek, H., 2014. Gender, competitiveness and career choices. *Quart. J. Econ.* 129, 1409–1447.
- Cardenas, J.-C., Dreber, A., von Essen, E., Ranehill, E., 2012. Gender differences in competitiveness and risk taking: comparing children in Colombia and Sweden. *J. Econ. Behav. Org.* 83, 11–23.
- Crosan, R., Gneezy, U., 2009. Gender differences in preferences. *J. Econ. Lit.* 47, 448–454.
- Dreber, A., von Essen, E., Ranehill, E., 2011. Outrunning the gender gap – boys and girls compete equally. *Exp. Econ.* 14, 567–582.
- Dreber, A., von Essen, E., Ranehill, E., 2014. Gender and competition in adolescence: task matters. *Exp. Econ.* 17, 154–172.
- Eckel, C.C., Grossman, P.J., 2008a. Differences in the economic decisions of men and women: experimental evidence. In: Plott, C.R., Smith, V.L. (Eds.), *Handbook of Experimental Economics Results*, vol. 1. Elsevier, pp. 509–519 (Chapter 57).
- Eckel, C.C., Grossman, P.J., 2008b. Men, women and risk aversion: experimental evidence. In: Plott, C.R., Smith, V.L. (Eds.), *Handbook of Experimental Economics Results*, vol. 1. Elsevier, pp. 1061–1073 (Chapter 113).

- Engel, C., 2011. Dictator games: a meta study. *Exp. Econ.* 14, 583–610.
- Fehr, E., Glatzle-Rutzler, D., Sutter, M., 2013. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *Eur. Econ. Rev.* 64, 369–383.
- Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: evidence from a matrilineal and a patriarchal society. *Econometrica* 77, 1637–1664.
- Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. *Quart. J. Econ.* 118, 1049–1074.
- Gneezy, U., Rustichini, A., 2004. Gender and competition at a young age. *Am. Econ. Rev.: Pap. Proc.* 94, 377–381.
- Grosse, N.D., Reiner, G., 2010. Explaining Gender Differences in Competitiveness: Gender-Task Stereotypes, Jena Economic Research Papers 2010-17.
- Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., Hummel, A., 2010. Preschoolers' allocations in the dictator game: the role of moral emotions. *J. Econ. Psychol.* 31, 25–34.
- Gunther, C., Ekinici, N.A., Schwieren, C., Strobel, M., 2009. Women can't jump? An experiment on competitive attitudes and stereotype threat. *J. Econ. Behav. Org.* 75, 395–401.
- Harbaugh, W.T., Krause, K., Vesterlund, L., 2002. Risk attitudes of children and adults: choices over small and large probability gains and losses. *Exp. Econ.* 5, 53–84.
- Hausmann, R., Tyson, L.D., Bekhouche, Y., Zahidi, S., 2013. The Global Gender Gap Report 2013, Tech. Rep. In: World Economic Forum, Geneva, Switzerland.
- Kamas, L., Preston, A., 2012. The importance of being confident; gender, career choice, and willingness to compete. *J. Econ. Behav. Org.* 83, 82–97.
- Kaser, K., 2008. *Gender Relations in Turkey and in the Balkans, 1500–2000*. LitVerlag, Munster.
- Michael, M., King, L., Guo, L., McKee, M., Richardson, E., Stuckler, D., 2013. The mystery of missing female children in the caucasus: an analysis of sex ratios by birth order. *Int. Persp. Sex. Reprod. Health* 39, 97–102.
- Mullis, I.V., Martin, M.O., Foy, P., Arora, A., 2012. TIMSS 2011 International Results in Mathematics. TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA), Boston, MA.
- Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? *Quart. J. Econ.* 122, 1067–1101.
- Niederle, M., Vesterlund, L., 2010. Explaining the gender gap in math test scores: the role of competition. *J. Econ. Persp.* 24, 129–144.
- Niederle, M., Yestrumskas, A.H., 2008. Gender Differences in Seeking Challenges: The Role of Institutions (unpublished).
- Samak, A.C., 2013. Is there a gender gap in preschoolers' competitiveness? An experiment in the U.S. *J. Econ. Behav. Org.* 92, 22–31.
- Sapienza, P., Zingales, L., Maestripieri, D., 2009. Gender differences in financial risk aversion and career choices are affected by testosterone. *Proc. Natl. Acad. Sci. U. S. A.* 106, 15268–15273.
- Shurchkov, O., 2012. Under pressure: gender differences in output quality and quantity under competition and time constraints. *J. Eur. Econ. Assoc.* 10, 1189–1213.
- Sutter, M., Glatzle-Rutzler, D., 2015. Gender differences in competition emerge early in life and persist. *Manage. Sci.* (in press).
- Sutter, M., Kocher, M.G., Rutzler, D., Trautmann, S.T., 2013. Impatience and uncertainty: experimental decisions predict adolescents' field behavior. *Am. Econ. Rev.* 103, 510–531.
- Wozniak, D., Harbaugh, W.T., Mayr, U., 2014. The menstrual cycle and performance feedback alter gender differences in competitive choices. *J. Labor Econ.* 32, 161–198.
- Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Hirschberg, A.L., Johannesson, M., 2009. A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proc. Natl. Acad. Sci. U. S. A.* 106, 6535–6538.
- Zhang, Y.J., 2013. Can Experimental Economics Explain Competitive Behavior Outside the Lab? (unpublished).
- Zhang, Y.J., 2014. Culture and the Gender Gap in Competitive Inclination: Evidence from the Communist Experiment in China (unpublished).