Gender Differences in Preferences at a Young Age? Experimental Evidence from Armenia*

Karen Khachatryan†
Stockholm School of Economics

June 2012

Abstract

We look at gender differences in competitiveness, risk preferences and altruism among a sample of 824 children and adolescents aged 8 to 16 in Armenia. Exploring four different competition tasks, girls are significantly more competitive in one task when it comes to performance change, and there are no gender differences in the other tasks or in the propensity to choose to compete. We find that girls are more altruistic and less risk taking than boys, and that the latter gap appears around the age of puberty. These results suggest that gender gaps in competitiveness are not always present.

Keywords: Gender differences, development of preferences, experiments with children, competitiveness, risk preferences, dictator game, altruism.

JEL codes: C91, C93, D03, J16, J33

*I would like to thank Anna Dreber, Emma von Essen, Tore Ellingsen, Magnus Johannesson, Eva Ranehill, Jörgen Weibull and seminar participants at the Stockholm School of Economics, University of Lancaster, University of Essex, and audiences at the Economic Science Association International Meeting in Chicago (July 2011), the 6th Nordic Conference on Behavioral and Experimental Economics in Lund (October 2011), Annual Congress of the Spanish Economic Association in Malaga (December 2011), and the Royal Economics Society Annual Conference in Cambridge (March 2012), for their useful comments, suggestions and discussions. Anna Dreber, Emma von Essen and Eva Ranehill sparked my interest in this topic, and without their help and encouragement I would have never started it. Moreover, Emma von Essen was kind enough to help me train the experimenters in Armenia, for which I am grateful. I also thank the Knut and Alice Wallenberg Research Foundation and the Foundation for Bank Research (Bankforskningsinstitutet) for financial support.

†Department of Economics, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm, Sweden. Email: karen.khachatryan@hhs.se; karen.g.khachatryan@gmail.com.

1. Introduction

Although women in many countries are as likely as men to pursue higher education and participate in the labor market, men still dominate top positions in most sectors in most societies. A number of reasons for this have been proposed in the economics literature including taste based or statistical discrimination, as well as gender differences in preferences, with the focus on competitiveness, risk preferences and social preferences. Women are in general found to be less risk taking than men, and sometimes also less competitive and more altruistic (see, e.g., Bertrand, 2010; Croson and Gneezy, 2009; Eckel and Grossman, 2008a,b; Engel, 2011).

Meanwhile, relatively little is known about the development of the gender gaps in economic preferences, to what extent children and adolescents exhibit the same type of gender differences in preferences as adults do, and to what extent culture and context might matter. This paper aims to contribute to further this understanding.

In this paper we explore the gender gap in preferences among children and adolescents in Armenia. We focus on competitiveness, risk preferences and altruism since these are the three areas in which gender differences are often found. Competitiveness is typically measured by either the performance change in response to a competitive setting compared to a noncompetitive setting, or the choice of whether to compete or not when given the choice between a competitive setting and a noncompetitive setting. We study competitiveness using both measures in running, skipping rope, math and word search, as well as competitiveness at the choice whether to compete or not in math and word search. We study risk preferences by having the subjects choose between different certain amounts and a gamble. Finally, we study altruism through a dictator game where the recipient is a charity.

We find that boys and girls are equally competitive when it comes to performance change in skipping rope, math and word search, whereas girls are more competitive than boys in running. The latter result is different from what has previously been found among children. However, this difference is only present among older children in our sample. There is no gender gap in the choice to compete or not in math or word search. We find that boys are more risk taking than girls and that this gap appears around the age of 12. We also find that girls are significantly more altruistic than boys.

When it comes to competitiveness, most previous related studies find that if there is a gender gap in any of the measures, men and boys tend to be more competitive than women and girls (e.g. Gneezy et al., 2003; Gneezy and Rustichini, 2004b; Gupta et al., forthcoming; Niederle and Vesterlund, 2007; Sutter and Rutzler, 2010). However, the gap can be influenced by the sample in which competitiveness is studied, concerning both country and age group. Among adults, the impact of culture on the gender gap in competitiveness has been shown by Gneezy et al. (2009), who find that men are more competitive than women in a patriarchal society in Tanzania whereas this gender gap is reversed in a matrilineal society in India. The task performed also seems to matter. In some cases (Gneezy and Rustichini, 2004a; Gunther et al., 2009; Kamas and Preston, 2010; Niels and Reiner, 2010; Shurchkov, forthcoming) but not others (Wozniak et al., 2011) the gender gap in competitiveness among adults diminishes when the task performed is word related compared to, for example, solving mazes or simple math.

Among children, tasks and culture have also been shown to influence the size and existence of a gender gap in competitiveness. However, the results are somewhat mixed (Andersen et al., 2011; Booth and Nolen, 2011a; Sutter and Rutzler, 2010; Zhang, 2011). For example, Dreber et al. (2011b) find no difference in the gender gap between running, skipping rope and dancing, whereas Cardenas et al. (forthcoming) find some influence of the task but only in Sweden. When it comes to cultures, the results are also mixed. For example, Gneezy and Rustichini (2004b) find that boys but not girls are competitive when it comes to performance change in running in Israel, whereas Dreber et al. (2011b) find no gender gap in Sweden and Cardenas et al. (forthcoming) find no gender gap in Colombia or Sweden with this measure in the same task.

Studies on risk taking among adults typically find that women are less risk taking than men (Croson and Gneezy, 2009). When it comes to children, Harbaugh et al. (2002) find no gender gap, whereas Borghans et al. (2009); Dreber et al. (2011b); Sutter et al. (2010) and Cardenas et al. (forthcoming) find that boys are more risk taking than girls. As for competitiveness, there is some evidence that the gender gap in risk taking seems also to be influenced by the context or sample studied. Booth and Nolen (2011b) compare children around the age of 15 in single sex and mixed schools and find that boys are more risk taking than girls in mixed schools but that there is no gender gap when comparing boys to girls from single sex schools. Girls are also more risk taking when assigned to all-girl groups than when assigned to mixed groups.

When it comes to altruism, Engel (2011) performs a meta-analysis of the experiments on the dictator game among adults and finds that women are more altruistic than men. The results from studies on children are thus far mixed, where some find that girls are more altruistic (Dreber et al., 2011a; Gummerum et al., 2010; Harbaugh et al., 2003), some that there is no gender gap (Benenson et al., 2007; Blake and Rand, 2010), and one study finds that girls are less altruistic (Fehr et al., 2011).[1]

Our results add to the literature on the importance of studying different samples, cultures and age groups when exploring the foundations of (gender differences in) economic preferences. Understanding the (development of) gender gaps in competitiveness, risk preferences and altruism will hopefully further our understanding of the gender gaps observed in important economic outcomes such as those related to the labor market (see, e.g., Bertrand, 2010, for further discussion). Armenia is different from most countries previously studied in that it was part of the Soviet Union and that perhaps the communist policies aimed at influencing the position of women in Armenia during the Soviet era had a long lasting and deep impact.

[1] Even though there is no gender gap in the amount given in the dictator game in Blake and Rand (2010), girls are more likely than boys to give something compared to nothing.

The rest of the paper, which reports the results of two large scale economic experiments, is organized as follows. We present the experimental setup and procedures in Section 2. Results for the competitiveness part of the study are presented in Section 3, for risk preferences in Section 4 and altruism in Section 5. We finish by a discussion in Section 6.

2. Experimental Design and Procedures

Our experiment was divided into two major parts. The experimental design and procedures closely follow to that of Cardenas et al. (forthcoming).[2] The experiments took place in two nearby secondary schools in the capital of Armenia, Yerevan, during a four week period in April–May 2010.[3] Overall 824 students aged 8 to 16 (in school years 2 to 10, 428 boys and 396 girls) participated in the study.

The first study took place during regular physical education classes with students in years 2 to 9 and used two experimental tasks: running and skipping/jumping rope.[4] Before the study began the subjects were told that they would participate in a series of tasks. In this part of the experiment, no extrinsic compensation was awarded and we relied on the intrinsic motivation that comes from winning a race or a competition, as in Gneezy and Rustichini (2004b) and Dreber et al. (2011b).

The subjects performed the tasks in a random order and in the presence of their classmates. Both the running and skipping rope tasks consisted of two rounds. In the first round, the subjects performed the task alone. In the second round the subjects performed the task in pairs. While performing the tasks in the first round the subjects were unaware of the existence of the second round. In the second round, after all the students had completed all first-round tasks, the students were matched with someone who had a similar performance to themselves in the first round. If more than two students obtained the same result in the first round, the matching was random. The students were made aware of the exact matching procedure. We also let a separate group of students perform the task alone in the second round to serve as control for unobservable characteristics that might differ between genders.

Performance in running was based on how fast the children ran 4 times 13 meters (for a total of 52 meters) and was the same for all the subjects. When competing in running in the second round, the students started at the same time and ran parallel to each other. The skipping rope task was different for younger and older children. In the skipping rope task, students in years 2 to 6 jumped with a long rope that one teacher or experimenter and one child turned and performance was measured by the number of jumps until the child first missed. Students in years 7 to 9 skipped

[2]In particular, for younger children the experimental design and procedures are identical. This was chosen in order to facilitate potential future comparison with the data gathered in Colombian and Sweden.

[3]To help run the experiments, we recruited and trained four young female experimenters. For the first study, they all worked together as a team overseeing different tasks, each specializing in one or two tasks, i.e., running times were always taken by the same experimenter. For the second study (classroom part), the experimenters worked in teams of two.

[4]The students also participated in a cooperation task, but this is outside the scope of the current paper.

a regular jumping rope for two minutes and performance was measured by the number of jumps in these two minutes. When competing in skipping rope in the second round, two ropes were put next to each other and the students were instructed to start jumping 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, the most common measure of the reaction to competition.

The second study took place during regular class hours, in their own classrooms, with students in years 2 to 10. At the start of the experiment the students were told that they are taking part in an economic experiment and can earn prizes or money by earning points through various tasks and that more points would correspond to more prizes or more money. As an extrinsic motivation the younger children — students in years 2 to 6 — were rewarded with pens, while older subjects — students in years 7 to 10 — were rewarded with money. Information was always revealed sequentially as the experiment progressed and at no point in time did the subjects get any feedback about their performance at any stage.

The students started with either a math or a word search task. Examples of these tasks can be found in Appendix A. The order of these tasks was randomized for each class and year. Each of these tasks consisted of three main stages. Performance in both tasks was measured as the number of correct answers: the number of correctly solved problems in the math task and the number of correct words found in the word search task. The students were sequentially informed of the incentive structure of each stage, which was as follows.

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

Stage 2: Tournament. Students were again asked to solve as many problems as possible in 2 minutes. 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. Students were asked to choose either Option One or Option Two and then solve as many problems as possible in 2 minutes. If a student chose Option One, she would get 3 points per solved problem. If she chose Option Two, she would be randomly paired with another student in her own class and she would get 6 points per correctly solved problem if she solved at least as many problems as the other student, otherwise she would get 0 points as in Stage 2.

Comparing performance in the second stage with performance in the first stage gives us a measure of competitiveness as absolute performance change or reaction to competition, whereas the choice in the third stage gives as a measure of competitiveness as a preference for competition. In the last stage of each task, Stage 4, we asked the children to guess how many children they believed had performed better than they did for both the first stage (piece-rate scheme) and the second stage (forced tournament).[5]

After the math and word search tasks the students participated in a risk task. The risk task consisted of six choices where the students could choose between a 50/50 gamble, a lottery in the form of a coin flip that gave 10 or 0 points and a safe option, where the certain amount increased successively in points, from 2 to 7.5. Our first measure of risk preferences relies on the unique switching point where the subject switches from preferring the gamble to preferring the safe option. This measure excludes inconsistent subjects, i.e. those with multiple switching points. Since some of our subjects are inconsistent in their choices, as in most other studies with similar measures, we also employ an alternative measure of risk preferences in order not to exclude the inconsistent subjects. This alternative measure is defined by the number of times an individual chooses a risky option, the gamble, over a safe option.

Next, the students took part in a dictator game, where they were asked to allocate 100 experimental points between themselves and a well known charity organization, Zeitoun Orphanage in Yerevan. The children were informed that the points they allocated to the charity would be converted to gifts and money and sent by the experimenters to the orphanage at the end of the study. The amount that a student donates to the charity is our measure of altruistic behavior.

Finally, at the end of the study, the students were administered an "exit" survey in order to measure different beliefs concerning the different tasks, motivation, cooperation, competition as well as collect some demographics.[6] After the experiment was over the experimenters collected and corrected all the papers and the students received their rewards.

Summarizing, the experimental design allows as to analyze (1a) competitiveness as performance change in running, skipping rope, math and word search, (1b) competitiveness as preference for competition in math and word search, (2) risk preferences through incentivized choices over lotteries and safe options, and (3) altruistic behavior via a dictator game.

3. Results: Competition

In this section we test whether there is a gender gap in competitiveness among children and adolescents in Armenia and whether the nature of the experimental task affects the size and direction of the gender gap. We start by looking at gender differences in competitive behavior in the running and skipping rope tasks. Then we address the effect of the gender composition in the competitive setting in these same tasks and present a robustness test. Thereafter we turn to looking at gender differences in competitive behavior in math and word search tasks.

[5] Since we do not observe differential gender selection into competitive environments, performance beliefs, or under and overconfidence, are not explored in this version of the paper.

[6] In this classroom part of the study the students also participated in a public goods game and a time preference task, but analysis of these data are outside the scope of the current paper and will appear in a future paper.

Throughout the paper all tests of the differences in means are analyzed using the nonparametric Wilcoxon-Mann-Whitney U test and a two sided t test. Only the $p-$values for the Mann-Whitney U tests are reported unless otherwise noted.[7] All the error bars in the graphs are 95% confidence intervals based on a normal distribution.[8]

3.1. Competitiveness in Running and Skipping Rope

For the analysis in this subsection and for descriptive purposes, we find it useful to divide the whole sample into three age groups, based on the year in school: years 2 to 4; years 5 to 6; and years 7 to 9.[9] We will refer to these samples as the younger, middle, and older samples, respectively. As noted before the skipping rope task was different for the older sample. This division allows us to compare performance distributions between genders and also across age more conveniently. Throughout the analysis, we also report results based on the whole sample.

3.1.1. Competitiveness in the Running Task

In the running and skipping rope tasks we had two groups of school children performing the task (competitive and control treatments). Since all conditions were identical in the first round, we can pool the outcomes to test for a gender difference in speed. Figure 1 presents a summary plot of the distribution of running times by gender and by age group is presented. We find that in the individual setting, in all the samples, boys ran on average faster than girls. The $p-$values for a significant gender difference are less than 0.001, see Table 1.[10] We note that this result is different from Gneezy and Rustichini (2004b) and similar to Dreber et al. (2011b).

When it comes to competitiveness, Table 2 shows that average performance in the competitive Round 2 differs significantly from performance in the noncompetitive Round 1 (Wilcoxon matched pairs signed-rank test, henceforth SR test in the tables). In all the samples both genders improve their performance significantly in running.

Figure 2 shows the distribution of the performance change in running for all the samples under study. In the whole sample, there are significant gender differences ($p < 0.001$): girls compete more. This result is mainly driven by older girls. In the younger sample of school children this gender difference is not significant ($p = 0.257$). The $p-$value for a significant gender difference in the middle and older samples of school children are 0.012 and 0.002, respectively.

---

[7]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.

[8]We should note that none of the relevant variables are normally distributed when employing standard tests of normality, e.g., Shapiro-Wilk and Skewness/Kurtosis tests.

[9]Note that students in year 10 did not participate in this part of the study.

[10]Additional regression analysis of time in the first round on age and a female dummy shows that girls ran on average 1.62 seconds slower compared to boys, and an additional year in age resulted in 0.49 seconds of improvement in speed.

Table 1 – Speed in the First Round

Age Group	Boys	Girls	All	M–W test (p–value)
Years 2 to 4	17.76 (1.68) 136	19.14 (1.60) 113	18.39 (1.78) 249	< 0.001
Years 5 to 6	16.80 (1.16) 63	18.47 (1.42) 69	17.68 (1.54) 132	< 0.001
Years 7 to 9	15.03 (1.50) 124	16.74 (2.08) 122	15.88 (2.01) 246	< 0.001
All	16.53 (1.96) 323	18.03 (2.08) 304	17.25 (2.15) 627	< 0.001

Notes: The table shows the means as the main number and the standard deviations in parentheses. Number of observations are below the means.

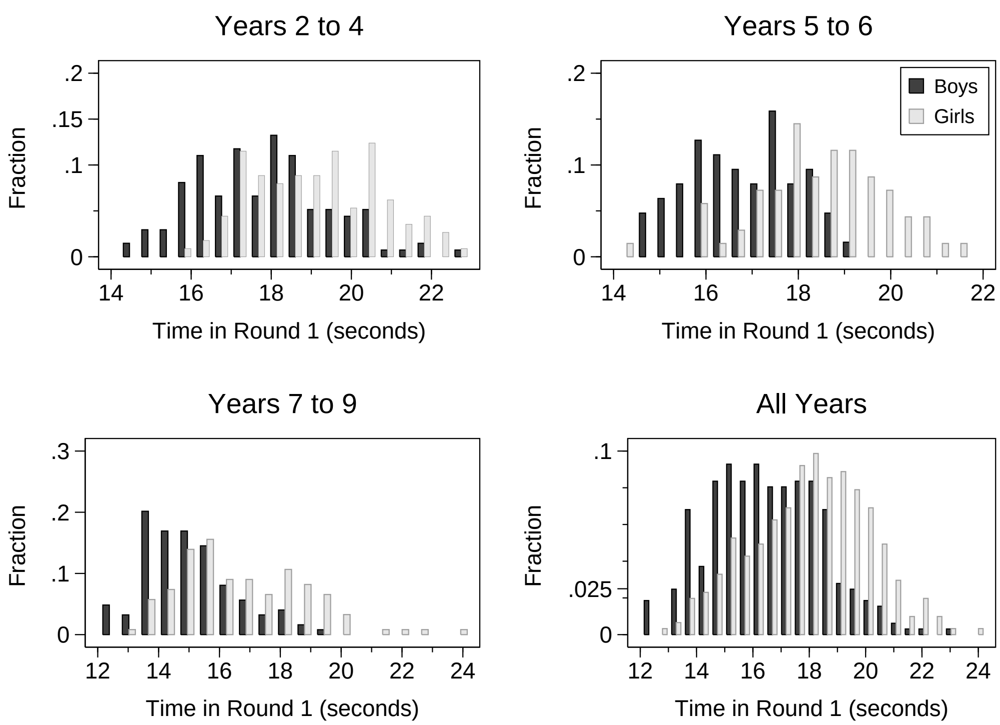

Figure 1. Distribution of running times in the first round, by gender and age group

Table 2 – Performance in the Running Task (seconds)

Age Group	Sex	Round 1	Round 2	SR test ($p-$value)	Number of observations
Years 2 to 4	Boys	17.67	17.14	< 0.001	100
	Girls	19.27	18.60	< 0.001	87
Years 5 to 6	Boys	16.89	16.67	0.005	54
	Girls	18.50	17.93	< 0.001	60
Years 7 to 9	Boys	15.13	14.92	0.001	100
	Girls	16.94	16.27	< 0.001	91
All	Boys	16.50	16.17	< 0.001	254
	Girls	18.19	17.54	< 0.001	238
	Total	17.32	16.83	< 0.001	492

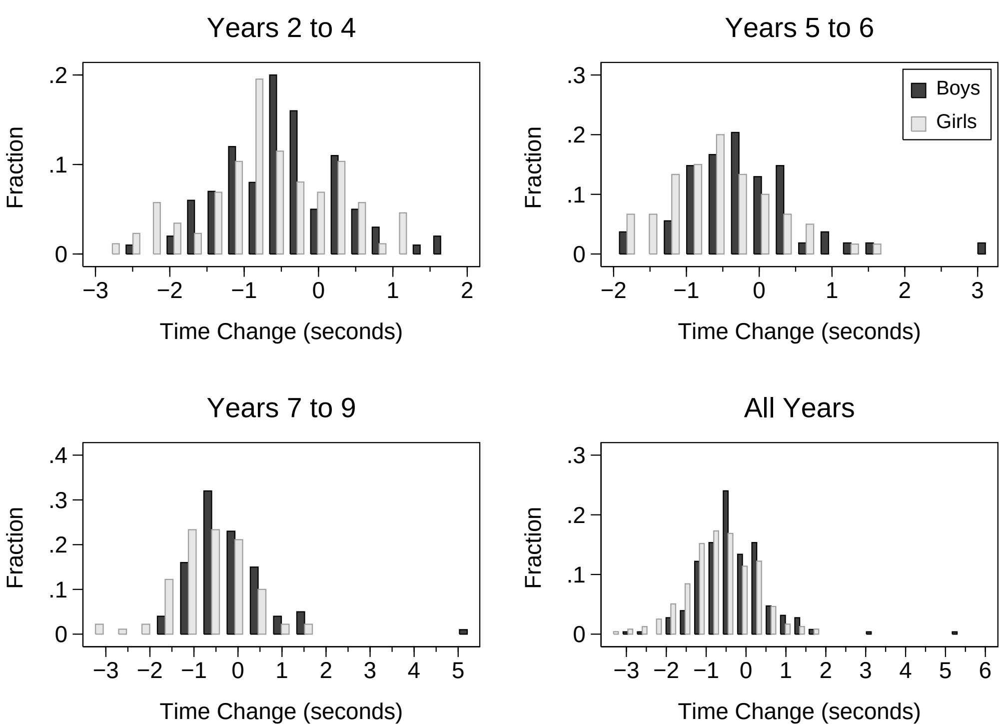

Figure 2. Distribution of running time changes, by gender and age group

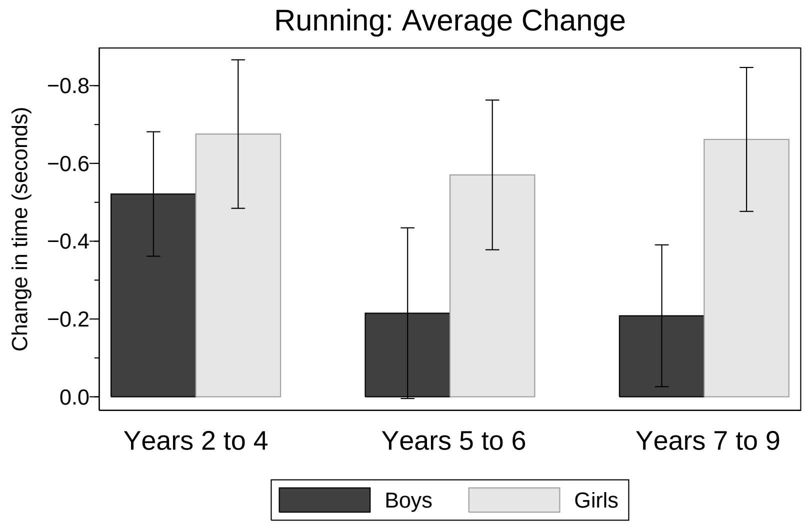

Figure 3. Average change in time, by gender and age group

Even though we find no significant gender difference in mean change in performance in the sample of younger children, there may be difference in the variances of the performance distributions. We test this and find no significant difference in the variance of change in performance between boys and girls in this sample. This is also true for all the other samples.[11]

The pattern for gender similarities and differences for different age groups are displayed in an aggregated manner in Figure 3. The plot shows the average change in performance by each gender within each sample. Girls on average improved 0.68 seconds in the younger sample, 0.57 seconds in the middle sample, 0.66 seconds in the older sample, and 0.64 seconds overall, or about 3.33%, 3%, 3.82%, and 3.43%, respectively. This can be compared to the average improvement for boys. Boys on average improved 0.52 seconds in the younger sample, 0.21 seconds in the middle and older samples, and 0.33 seconds overall, or about 2.86%, 1.26%, 1.3%, and 1.91%, respectively. As stated, the difference in average change in performance between boys and girls is not statistically significant only in the younger sample. The results are qualitatively the same for average relative change in performance.[12]

3.1.2. Competitiveness in the Skipping Rope Task

Since all the conditions are again identical in the first round we can pool the data to test for a gender difference in skipping rope. A summary plot of the distribution of number of jumps by gender and

[11] Since the most common test for comparison of standard deviations, the $F-$test for homogeneity of variances, relies on the assumption that the data are drawn from an underlying normal distribution and since none of our relevant variables are normally distributed, we also performed a robust test (Levene's test with mean, median and 10% trimmed mean). None of these tests indicated significant difference in the variances of the distributions.

[12] Average relative change in performance is measured as ((performance in round 2 – performance in round 1)/performance in round 1).

Table 3 – Number of Jumps in the First Round

Age Group	Boys	Girls	All	M–W test ($p-$value)
Years 2 to 4	5.31 (6.23) 135	6.17 (7.87) 110	5.70 (7.01) 245	0.565
Years 5 to 6	7.24 (9.39) 63	14.64 (17.30) 69	11.11 (14.52) 132	0.013
Years 7 to 9	64.96 (42.93) 119	71.17(28.69) 122	68.10 (36.48) 241	0.087
All	28.09 (39.27) 317	34.46 (36.80) 301	31.19 (38.19) 618	0.007

Notes: The table shows the means as the main number and the standard deviations in parentheses. Number of observations are below the means.

for each sample is presented in Figure 4. We find that in the individual setting, girls skip rope better than boys do, like in Dreber et al. (2011b). The $p-$value for a significant gender difference is 0.007, see Table 3. However, this gender difference is not significant in our sample of younger school children. It turned out that the younger children were not familiar with the task and had difficulty to perform it. Skipping rope used to be a common activity for young children in Armenia, but this is no longer true. Also, remember that the skipping rope task was not exactly the same for older and younger children, hence the much higher performance of the older children. When it comes to competitiveness, Table 4 shows that average performance in the competitive setting (Round 2) differs significantly from average performance in the noncompetitive setting (Round 1) only in the older sample. Again, this could be due to the fact that the younger children had difficulty to perform an unfamiliar task.

Figure 5 shows the distribution of performance change in skipping rope for all the three age groups and overall. There are no significant gender differences in the average change in performance in any of the samples under consideration ($p = 0.941$, $p = 0.724$, and $p = 0.527$, respectively). However, the variance of change in performance is significantly higher for girls in the younger sample (Levene's test, $p = 0.033$ at the median) and older sample (Levene's test, $p = 0.086$ at the median) but this difference also vanishes when we look at the variance of relative change in performance.

The pattern for gender similarities for different year groups are displayed in an aggregated manner in Figure 6. The plot shows the average change in performance by each gender within each sample. Girls on average improved by 1.48 jumps in the younger sample, deteriorated by 5.15 jumps

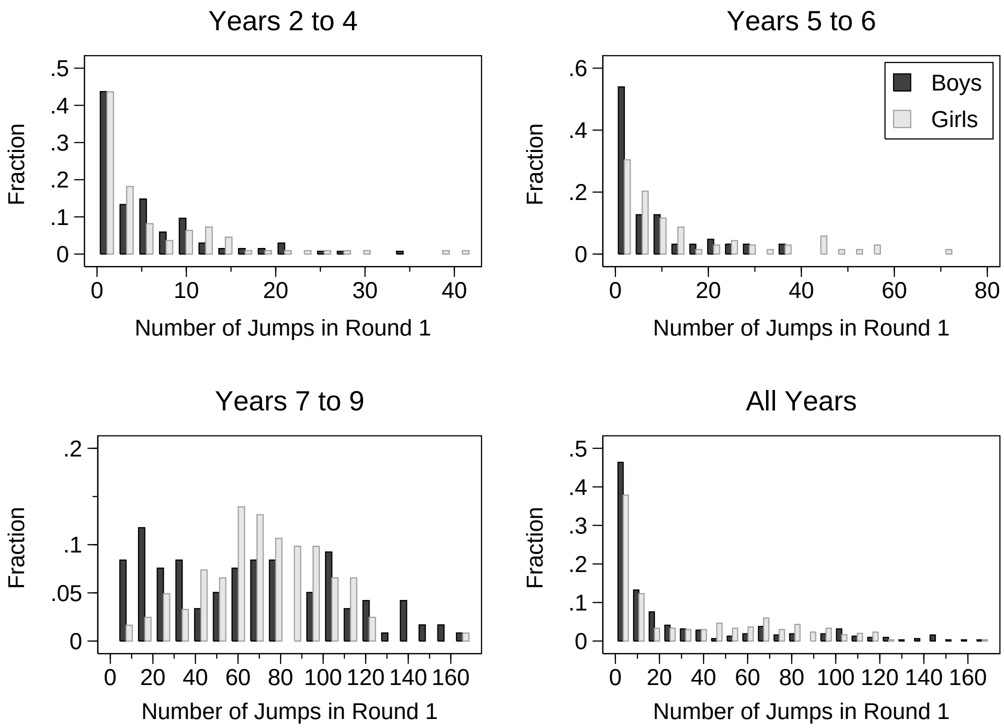

Figure 4. Distribution of jumps in the first round

Table 4 – Performance in the Skipping Rope Task

Age Group	Sex	Round 1	Round 2	SR test ($p-$value)	Number of observations
Years 2 to 4	Boys	5.14	6.52	0.084	99
	Girls	4.68	6.16	0.126	85
Years 5 to 6	Boys	7.35	6.17	0.539	54
	Girls	13.62	8.47	0.255	60
Years 7 to 9	Boys	67.10	72.27	< 0.001	96
	Girls	70.92	79.45	< 0.001	88
All	Boys	29.66	31.79	< 0.001	249
	Girls	32.50	34.43	< 0.001	233
	Total	31.04	33.07	< 0.001	482

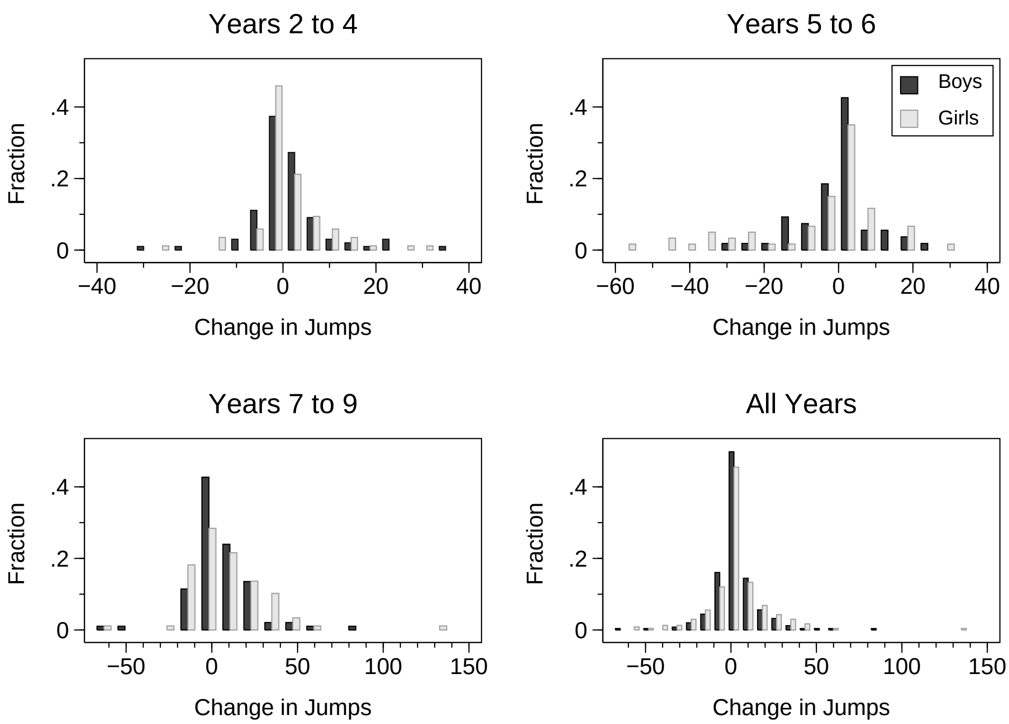

Figure 5. Distribution of changes in jumps, by gender and age group

in the middle sample, and improved by 8.93 jumps in the older sample, or about 111%, 49%, and 22.6% respectively.[13] This can be compared to the average improvement for boys. Boys on average improved by 1.37 jumps in the younger sample, deteriorated by 5.15 jumps in the middle sample, and improved by 5.51 jumps in the older sample, or about 107%, 113%, and 18.5%, respectively. As stated, the difference in average and relative changes in performance between boys and girls is not statistically significant in any of the three samples.

3.1.3. Impact of Opponent Gender on Competitive Behavior

Since the children could observe each other while performing in the second round, the gender of the opponent is known in both running and skipping rope. We find that both boys and girls compete significantly more against girls with the exception of the younger sample, in which we found no gender difference in reaction to competition. Table 5 gives an overall summary for the whole sample for the different pair compositions in our study for the running task and Table 6 presents the same information for the skipping rope task.

In running, both boys and girls improve the most when competing against a girl. This is also true for any of the samples under investigation. This difference in competitive behavior when

[13]Note, that this relative improvement measure ignores those subjects who made 0 jumps in the first round. 31 boys out of 198 and 23 girls out of 176 in years 2 to 6 performed 0 jumps.

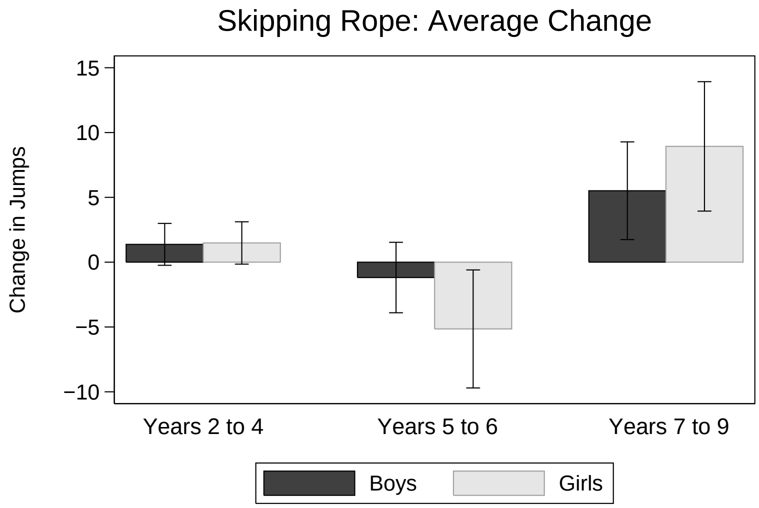

Figure 6. Average change in jumps, by gender and age group

Table 5 – Running: Performance Change Based on the Gender Composition of the Competing Pairs

Sample	Number of Obs.	Change in Time	Standard Error
Total	490	-0.484	0.039
Total Boys	253	-0.335	0.054
Total Girls	237	-0.644	0.055
Boys with boys	162	-0.248	0.066
Girls with girls	143	-0.764	0.073
Boys in mixed pairs	91	-0.489	0.092
Girls in mixed pairs	94	-0.461	0.083

Table 6 – Skipping Rope: Performance Change Based on the Gender Composition of the Competing Pairs

Sample	Number of Obs.	Change in Jumps	Standard Error
Total	479	2.499	0.739
Total Boys	246	2.415	1.340
Total Girls	233	2.588	1.206
Boys with boys	129	2.116	1.090
Girls with girls	116	2.673	1.371
Boys in mixed pairs	117	2.744	1.340
Girls in mixed pairs	117	2.603	1.996

facing the same vs. the opposite gender is statistically significant for both girls ($p = 0.006$) and boys ($p = 0.008$). However, as before, these significant differences are driven mainly by the older school children. In the younger and middle samples, the difference in competitive behavior when facing the same vs the opposite gender is statistically insignificant for both boys and girls ($p = 0.389$ and $p = 0.073$ for boys and girls, respectively in the combined sample of children in years 2 to 6).

In skipping rope, again both boys and girls improve the most when competing against a girl, but none of these differences is statistically significant in any of the samples for either gender.

Note that the extant research on the opponent gender effect is mixed. For example, among adults Gupta et al. (forthcoming) and Gneezy et al. (2003) find that women compete more against women and men more against men. Among children, on the contrary, Gneezy and Rustichini (2004b) find that boys are not affected by the gender composition of the competing pairs but girls compete more against boys, while Dreber et al. (2011b) find that neither boys nor girls are influenced by the gender of their opponent.

3.1.4. Robustness Check for the Running and Skipping Rope

As stated in the beginning of this section, we also let a separate group of children perform the task alone in the second stage, serving as a control group. We thereby control for unobservable factors that could cause differences in the outcome, such as one gender exerting more effort initially and getting tired faster than the other. The control group includes 135 subjects in the running task (69 boys and 66 girls) and 132 children in the skipping rope task (67 boys and 65 girls).

Children performing the tasks alone in the first and second rounds showed, on average, either no improvement or a slight improvement in the second round, depending on the task and the sample under investigation. Most importantly though, the difference in change of performance between genders in not significant in any of the tasks or samples.

We end this subsection by summarizing the results on competitiveness in running and skipping rope. When measuring competitiveness as a performance reaction to a competitive setting we found that girls compete more in running, but there is no gender gap in skipping rope. We also found significant opponent effects in running: both girls and boys improve their performance the most when competing against a girl. However, we should also note that the gender gap in running for the younger children was not significant.

3.2. Competitiveness in Math and Word Search

For the analysis in this subsection, we find it useful to divide the whole sample into three different age groups, based on the year in school: years 2 to 5; years 6 to 7; and years 8 to 10. We will refer to these samples as the younger, middle, and older samples, respectively. Note that this division differs slightly from the division in the previous subsection. This division is based on the premise that the level of the difficulty of the math task, and hence performance, was the same within each

Table 7 – Performance in the Math Task

Age Group	Sex	Stage 1 (piece rate)	Stage 2 (tournament)	SR test ($p-$value)	Number of observations
Years 2 to 5	Boys	19.64	22.06	< 0.001	177
	Girls	18.50	21.75	< 0.001	156
Years 6 to 7	Boys	11.19	15.88	< 0.001	95
	Girls	12.35	17.24	< 0.001	100
Years 8 to 10	Boys	6.03	7.46	< 0.001	116
	Girls	4.91	6.39	< 0.001	121
All	Boys	13.50	16.18	0.018	388
	Girls	12.50	15.62	0.001	377

of the three samples. Therefore comparing and interpreting differences in performance between genders is more straightforward in this manner. Throughout the analysis, we also report results based on the whole sample.

3.2.1. Competitiveness in Math and Word Search: Performance Change

Table 7 shows the average performance of boys and girls in Stage 1 (piece rate) and Stage 2 (compulsory tournament) of the math task. Generally, we see that boys are slightly better in the math task than girls. However, this gender difference in performance is not statistically significant in both stages ($p-$values for the first stage are 0.311, 0.298, 0.151, and 0.102 for the younger, middle, older, and the whole sample, respectively).[14]

In all the samples, both boys and girls are competitive in math in terms of reacting to competition (Wilcoxon matched-pairs signed rank tests, see Table 7). The increase from Stage 1 to Stage 2 is perhaps not only due to the higher incentives under the tournament scheme, but also to some learning going on. Figure 7 shows the distribution of performance change in math for boys and girls in all the samples. The figure shows that there are no gender differences in any of the distributions ($p-$values are 0.158, 0.703, 0.725, and 0.195 for the younger, middle, older samples, and the whole sample, respectively).[15] The pattern for gender similarities for different age groups are displayed in an aggregated manner in Figure 8. The plot shows the average change in performance by each

[14] When we look at gender differences in Stage 1 performance within each year, we find some evidence that boys in year 10 are significantly better than girls in math. However, this sample size is quite small (41 subjects: 24 girls and 17 boys).

[15] In the sample of older students boys have a significantly higher variance in performance change using the robust test, but the difference vanishes when we look at the distribution of relative performance change. Also boys in year 10 appear to be significantly more competitive than girls.

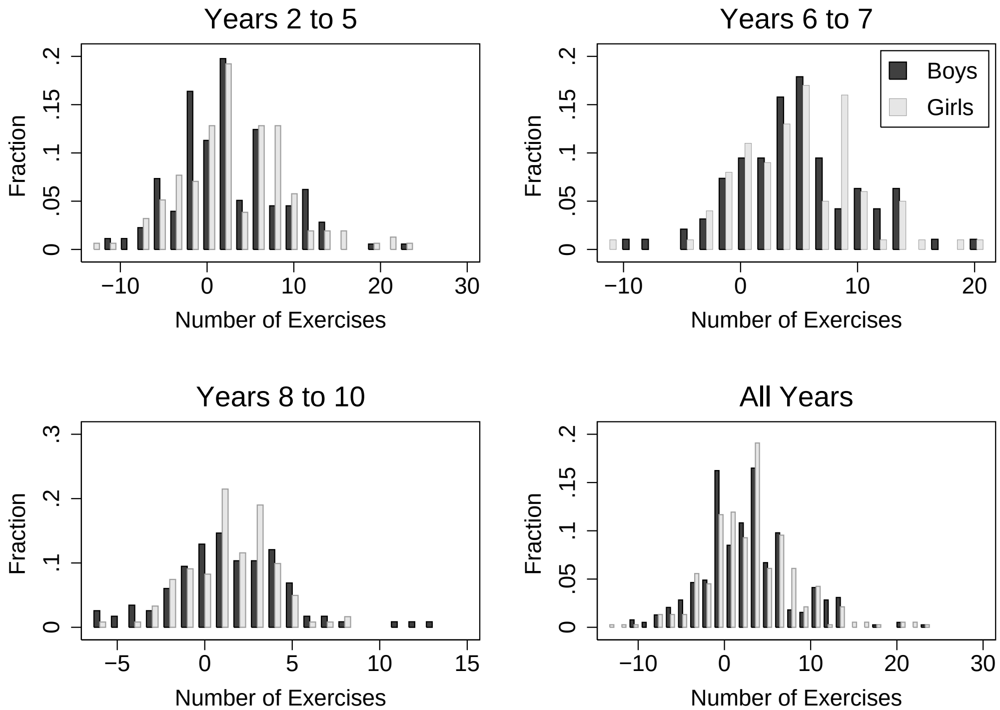

Figure 7. Distribution of performance change in the math task

gender within each age group.

Table 8 shows the average performance of boys and girls in Stage 1 (piece rate) and Stage 2 (compulsory tournament) of the word search task. Generally, we see that girls are better in word search than boys. This gender difference in performance is statistically significant in both stages for the whole sample ($p = 0.001$ and $p = 0.003$ in the first and second stages, respectively).[16]

When it comes to reacting to competition both boys and girls are competitive in word search (Wilcoxon matched-pairs signed rank tests, see Table 8). Figure 9 shows the distribution of performance change in word search for boys and girls in all the samples. The histograms show that there are no gender differences in any of the distributions ($p-$ values are $0.418, 0.716, 0.415$, and 0.324 for the younger, middle, older samples and the whole sample, respectively). This result is also robust to using a relative performance change as the measure of competitiveness. The pattern for gender similarities for different age groups are displayed in an aggregated manner in Figure 10. The plot shows the average change in performance by each gender within each age group in the word search task.

[16]When we look at gender differences in Stage 1 performance within each year, we find that there is no gender gap in ability in word search only in year 10. However, we should keep in mind that this sample is quite small (41 subjects: 24 girls and 17 boys).

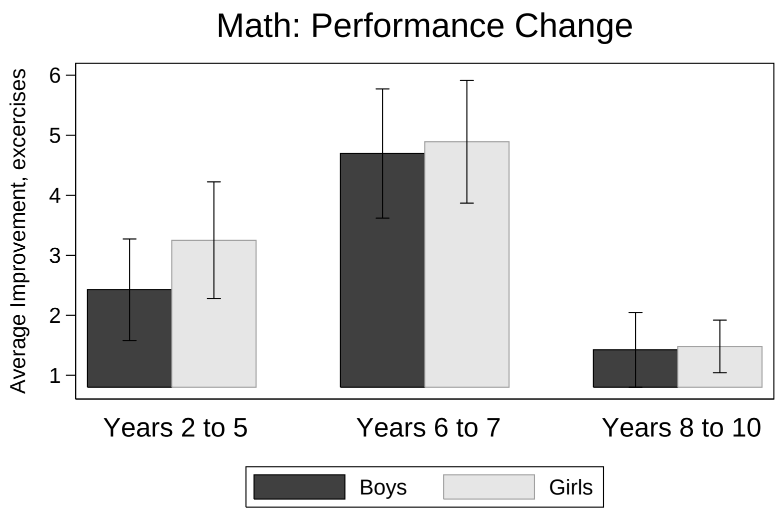

Figure 8. Average change in math performance, by gender and age group

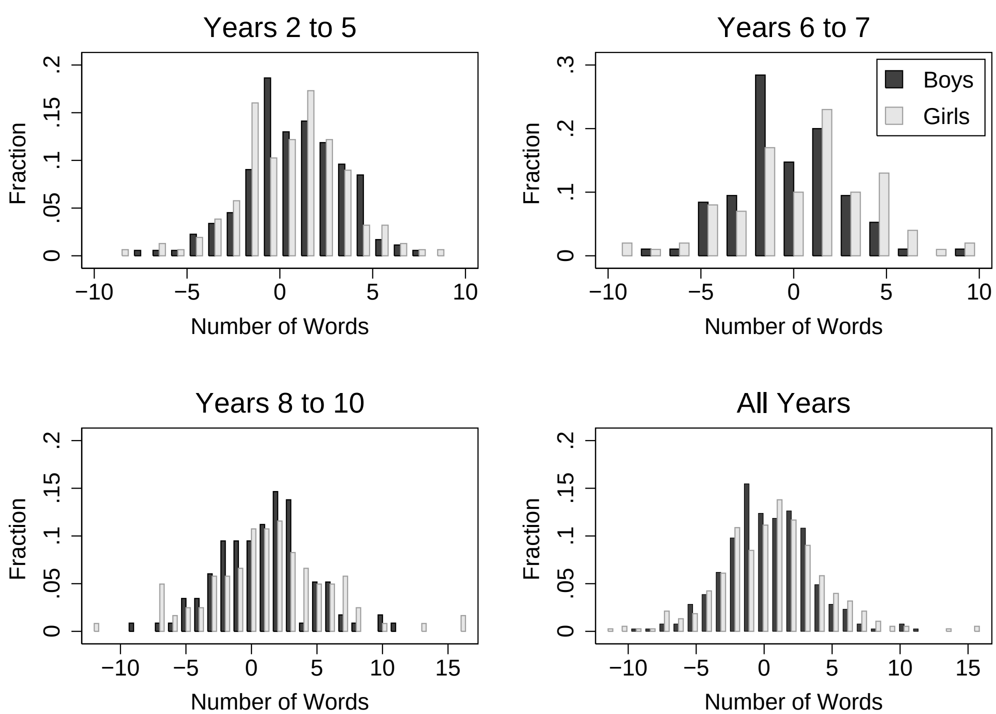

Figure 9. Distribution of performance change in the word search task

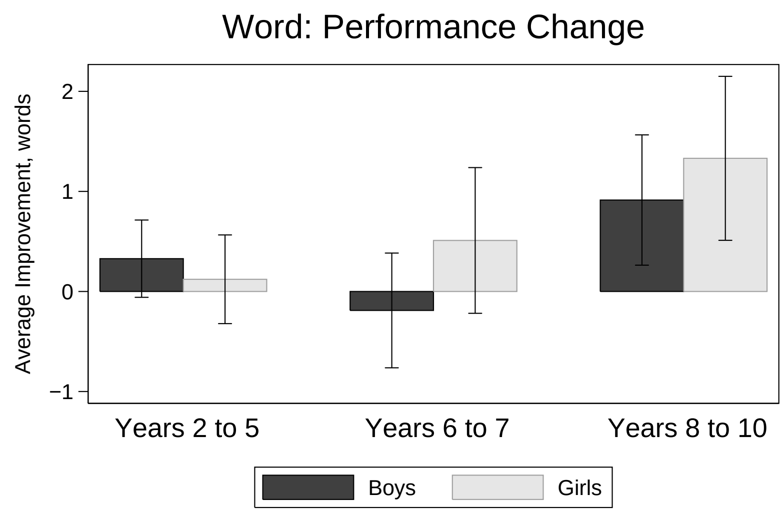

Figure 10. Average change in word performance, by gender and age group

Table 8 – Performance in the Word Search Task

Age Group	Sex	Stage 1 (piece rate)	Stage 2 (tournament)	SR test ($p-$value)	Number of observations
Years 2 to 5	Boys	4.91	5.24	0.068	177
	Girls	5.51	5.63	0.542	156
Years 6 to 7	Boys	7.31	7.12	0.385	95
	Girls	8.65	9.16	0.109	100
Years 8 to 10	Boys	9.78	10.69	0.008	116
	Girls	10.70	12.03	0.001	121
All	Boys	6.95	7.33	0.018	388
	Girls	8.01	8.62	0.001	377

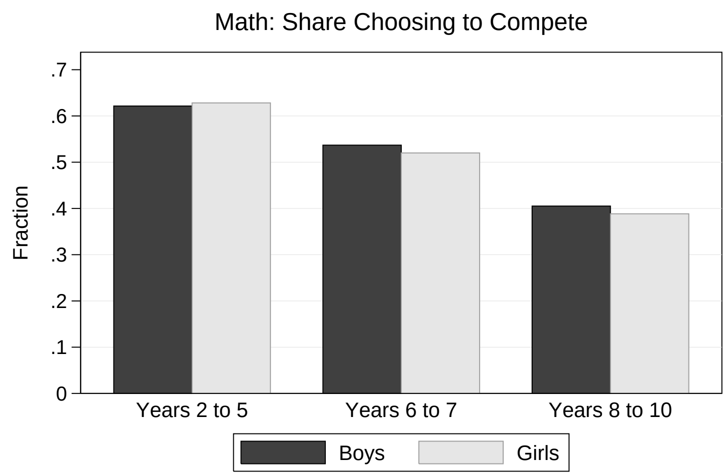

Figure 11. Math: Share choosing to compete

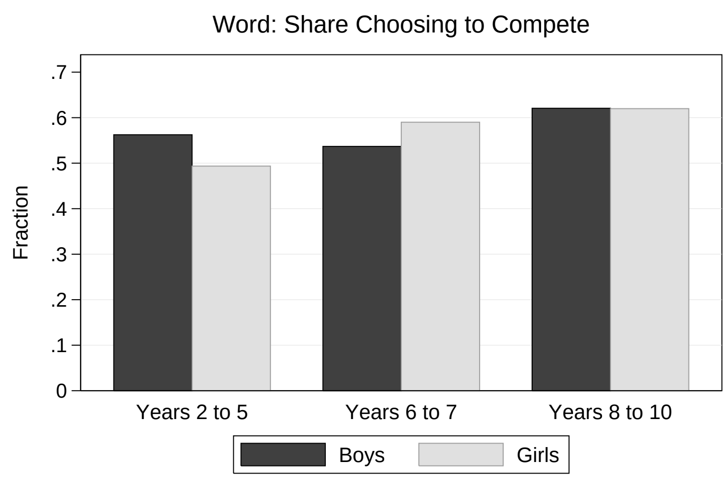

Figure 12. Word: Share choosing to compete

3.2.2. Competitiveness in Math and Word Search: Choice

In Stage 3 of math and word search tasks we let the subjects choose their compensation scheme. We find that boys and girls are equally likely to choose to compete in math and word search (math: $p = 0.497$, word: $p = 0.548$; one sided Fisher's exact test for the whole sample). 53.61% of boys and 52.25% of girls chose to compete in math, while 57.36% of boys and 55.97% of girls chose to compete in word search. Figures 11 and 12 shows the entry rates of girls and boys into the tournament payment scheme in Stage 3 of the experiment for each of the age groups. Within any of the samples there are no gender differences in competitiveness in terms of choice.

The only pattern that emerges is that in the math task younger children chose to compete significantly more than older ones. 62.46% of children in the younger sample chose to compete, compared with 52.82% in the middle, and 39.66% in the older samples respectively. These differences between each age group, and for each gender between each age group are very significant using a χ^2 test. These differences could be explained by the difficulty of the task. The level of the math task was significantly more difficult for the older sample than for the younger sample (see Appendix A).[17]

To summarize this subsection on competitiveness in math and word search, we found no gender gaps in competitiveness measured as either performance change in reaction to competition or by preference to enter into a competitive environment in either task.

3.3. Additional Analysis

At the end of the experiment we administered an "exit" survey, in which, among other things, we elicited perceptions of how boyish/girlish the subjects considered running, skipping rope, math and word search 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 subjects considered competing against a boy and against a girl to be (with 0 being not at all important and 10 being very important).

3.3.1. Do Subjects Perceive Competing to be Important?

On average, boys rate competing as more important than girls. However, this gender difference is not statistically significant ($p = 0.231$).[18] Boys believe it is more important to compete against a boy than against a girl ($p < 0.001$ in all age groups and overall). Girls believe it is more important to compete against a boy than against a girl, but the difference in ratings is not significant ($p = 0.079$). Interestingly enough, this does not correspond to what we observe in actual performance change in running and skipping rope. In both tasks, both boys and girls change their performance more when competing against a girl, see Tables 5 and 6.

3.3.2. Do Subjects Perceive the Tasks to be Gendered?

Boys perceive running to be significantly more boyish than girls do ($p < 0.001$), but the gap in ratings narrows with age. Boys and girls perceive skipping rope differently. Boys rate skipping rope as more gender neutral, while girls consider it to be more girlish and the difference in ratings is statistically significant ($p < 0.001$). When it comes to math boys consider it to be more boyish,

[17] Additional regression analysis for choice in math and word search tasks reveals no significant gender differences. The only significant variables affecting probability of choosing to compete are ability and risk aversion, and also age in the math task only.

[18] However, in the sample of students in years 8 to 10 there is some evidence of significant gender difference: $p = 0.026$ for a t-test, $p = 0.059$ for Mann-Whitney U test, and $p = 0.850$ for Kolmogorov-Smirnov test.

while girls consider it to be more gender neutral ($p < 0.001$), and the gap in rating somewhat narrows with age. When it comes to word search, both boys and girls perceive it to be closer to gender neutral but still differ in their opinions ($p < 0.001$).

4. Results: Risk Preferences

In this section we test whether there are gender differences in risk preferences among children and adolescents in Armenia. We look at the gender gap in risk preferences measured from incentivized lotteries in the classroom.

For the analysis in this section (and the next), we again find it useful to divide the subjects into roughly three equally sized groups based on age and year in school: years 2 to 4 (average age just under 9 years old); years 5 to 7 (average age of around 12 years old); and years 8 to 10 (average age just under 15 years old). We will refer to these groups as the younger, middle, and older samples, respectively.

Our main measure of risk taking relies on the unique switching point when subjects switch from preferring the safe option to a lottery. We measure risk taking as the certainty equivalent at the switching point, and take it to be the midpoint of the switching interval. In our sample of 762 schoolchildren, 18.63% of the subjects are inconsistent in their choices of the safe option versus the lottery (coin flip). However, most importantly, there is no gender difference in being inconsistent in the whole sample ($p = 0.274$, equality of proportions test) or any of the subsamples. There is some evidence though that the share of inconsistent subjects is disproportionately higher among younger children. Therefore, we also measure risk preferences in terms of the number of risky options chosen, in order not to exclude those subjects with inconsistent choices. Using this outcome measure, the results are qualitatively similar to those presented here and are relegated to an appendix, see Appendix B.

Table 9 presents summary statistics for the risk measures. In the aggregate, we find significant risk aversion in our sample, with a mean (median) measure of risk of 4.36 (4.5) ($p < 0.001$, Wilcoxon signed-rank test, testing whether it is different from 5). We find that boys take 16 percent more risk than girls do ($p < 0.001$). Figure 13 shows the raw distribution of sure amounts at which the subjects started to prefer the safe option by each gender for each of the samples.[19] The distribution of certainty equivalents will look visually the same. The histograms reveal that the distributions for risk preferences look markedly different for different age groups. The distribution for the older sample is remarkably different from the distribution of the younger sample.[20]

We find a gender gap in risk taking in all the age groups, with boys being more risk taking. However, the information presented in Table 10 and the histograms in Figure 13 clearly demonstrate

[19] The sure amount is coded as 10 for the subjects who always chose the risky option, and as 2 for the subjects who always chose the safe option.

[20] Perhaps younger children have difficulty in processing information about chance, or simply have different risk preferences.

Table 9 – Summary of Risk Measures

Variable	Sample	Mean	St. Dev.	Median	Min	Max	N
Risk	Boys	4.75	2.62	4.5	1	8.75	309
	Girls	3.97	2.36	3.5	1	8.75	311
	Total	4.36	2.52	4.5	1	8.75	620
Inconsistent	Boys	0.20	0.40	0	0	1	387
	Girls	0.17	0.38	0	0	1	375
	Total	0.19	0.39	0	0	1	762
Risky choices	Boys	3.02	1.94	3	0	6	387
	Girls	2.50	1.81	2	0	6	375
	Total	2.76	1.89	3	0	6	762

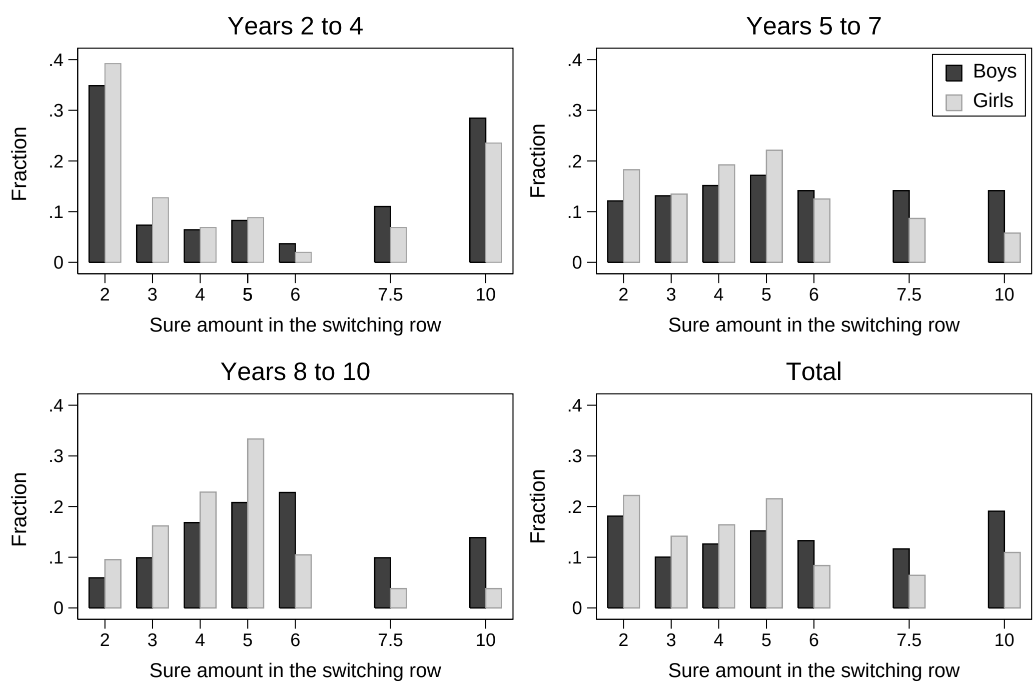

Figure 13. Distribution of risk preferences

Table 10 – Summary of Risk Taking

Age Group	Boys	Girls	All	M–W test ($p-$value)	K–S test ($p-$value)	t test ($p-$value)
Years 2 to 4	4.56 (0.31) 109	3.98 (0.31) 102	4.27 (0.22) 211	0.230	0.512	0.184
Years 5 to 7	4.72 (0.24) 99	3.96 (0.20) 104	4.33 (0.16) 203	0.026	0.137	0.016
Years 8 to 10	4.97 (0.21) 101	3.97 (0.16) 105	4.46 (0.14) 206	< 0.001	< 0.001	< 0.001
All	4.75 (0.15) 309	3.97 (0.13) 311	4.36 (0.10) 620	< 0.001	< 0.001	$< .0001$

Notes: The table shows the mean as the main number and the standard deviation in parentheses. Number of observations are below the means.

that the gender gap in risk taking is not statistically significant in the younger sample, and that we cannot reject the hypothesis that the distributions of risk preferences for boys and girls are the same in this age group (mean age 8.83). On the other hand, while there is mixed evidence that the gender gap in the middle sample is statistically significant, in the sample of older children (mean age 14.53) this gender gap is already highly significant. Older boys are much more risk taking than older girls. Figure 14 presents the gender gaps in an aggregated manner for the three samples.[21] We can see from the figure, that while on average risk preferences of girls does not change, there is almost a linear increase in average risk taking for boys with age. Comparing older children with younger ones we find that younger children take less risk than older children. This result is driven by the difference between older and younger boys, since girls are equally risk taking in the two groups.

In sum, we find that boys are more risk taking than girls and that this gender gap gets bigger during adolescence.

5. Results: Altruism

In this section we look at gender differences in altruism as measured via donations in a dictator game. We find that girls are significantly more altruistic than boys in our sample of students

[21] Doing the analysis for each age separately, shows that a significant gender gap appears at the age of 12 in our sample. Additional regression analysis of risk taking on being female and age shows a significant gender effect. When comparing gender gaps in risk taking between older and younger children in a regression framework, we find that the gap is significant only at the 10% level.

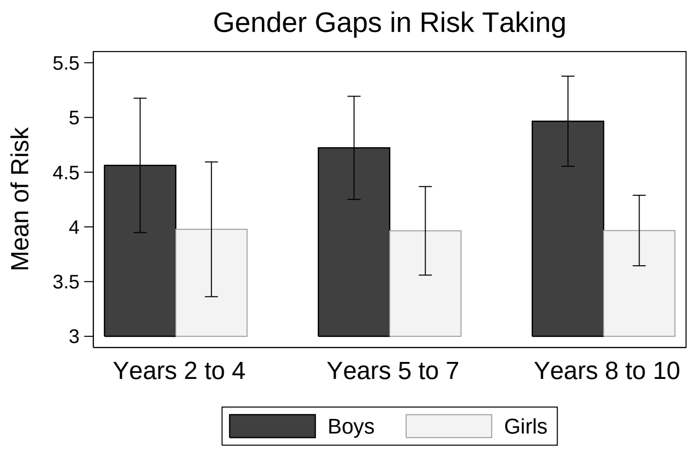

Figure 14. Average risk taking, by gender and age group

Table 11 – Donations in the Dictator Game

Sex	Number of observations	Average donation	Standard deviation
Boys	387	51.48	33.98
Girls	376	60.16	32.21
Total	763	55.75	33.38

($p < 0.001$). Girls on average donate 60 experimental points and boys 51 points out of 100 to the charity organization (an orphanage) that is the recipient in our dictator game (see Table 11 for a summary). The distribution of donations by gender is presented in Figure 15. The modal allocation is the 50–50 split.[22]

Upon a closer look we find that the gender gap is mainly driven by the behavior of older and younger boys. In the younger and older samples, the gender gap is significant ($p = 0.002$ and $p = 0.013$, respectively), while in the middle sample the gender gap is not significant ($p = 0.216$), see Figure 16.

Figure 17 shows the average donations for boys and girls in different years in school. The data show that there are clear gender differences in altruism for older and younger students, and this is mainly due to the behavior of the boys.

[22]Regression analysis controlling for age confirms a strong gender effect, although the coefficient of age is insignificant.

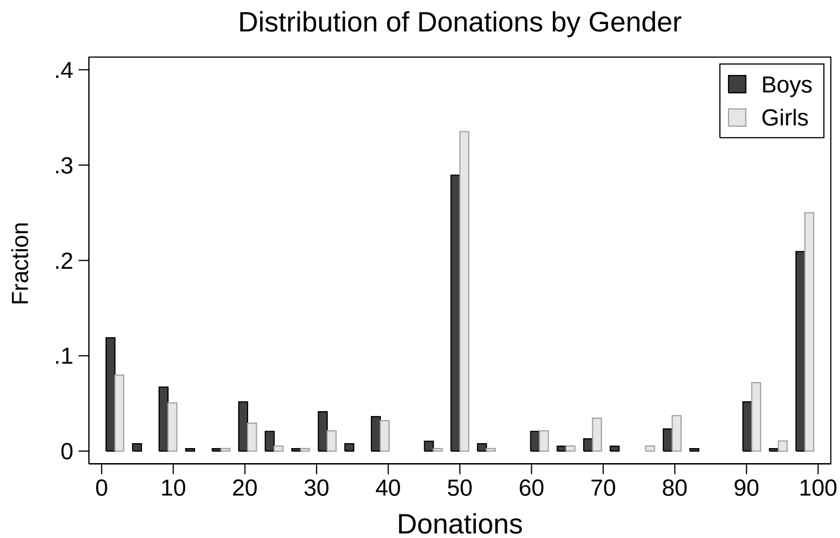

Figure 15. Distribution of donations in the dictator game, by gender

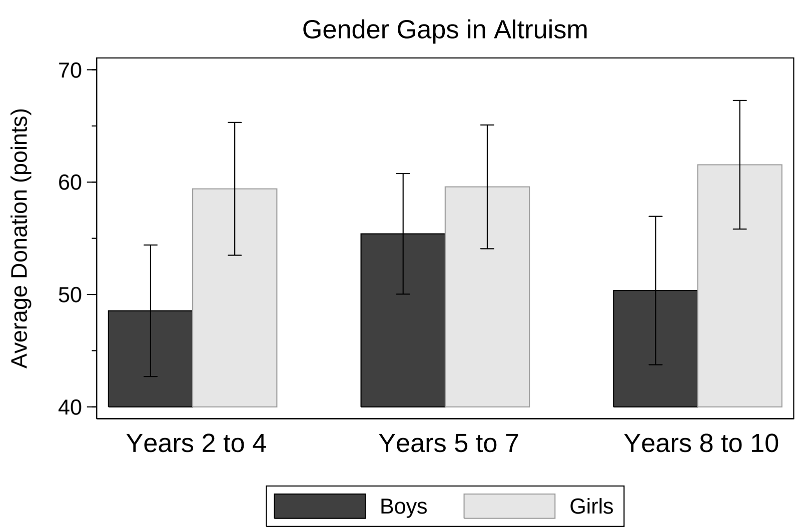

Figure 16. Average donations, by gender and age group

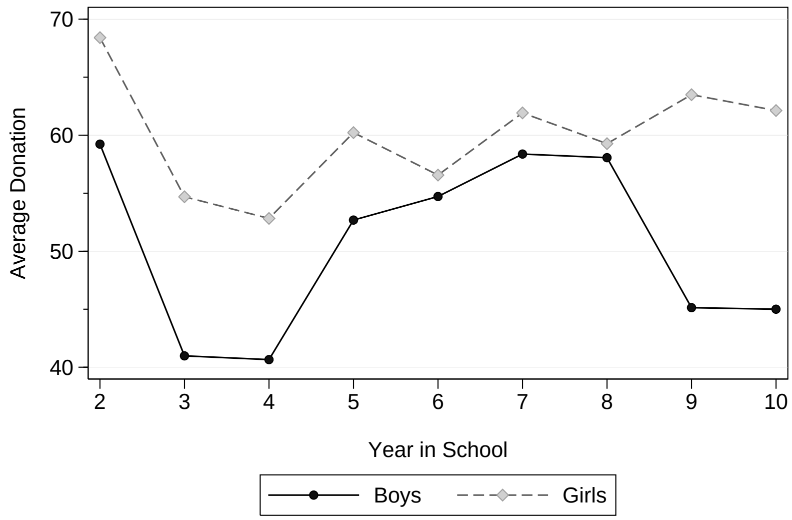

Figure 17. Average donation, by gender and year in school

6. Discussion and Conclusion

Recent papers have explored to what extent gender differences vary across cultures, contexts and age groups. There is some evidence that culture matters, but not always in predictable ways. For example, Gneezy and Rustichini (2004b) find that among children in Israel, only boys run faster when competing against another child compared to running alone, whereas Dreber et al. (2011b) find no gender gap in competitiveness with this measure in Sweden. Israel typically scores lower on gender equality indices than Sweden, which typically scores in the global top five. Another example that suggests that culture and norms, for example, about gender equality might matter in order to explain gender differences in preferences is Andersen et al. (2011). Comparing competitiveness in terms of the choice to compete or not among children aged 7 to 15 in a matrilineal society and patriarchal society in India, they find no gender gap in the matrilineal society, but that boys become more competitive than girls in the age group 13 to 15. However, Cardenas et al. (forthcoming) find gender differences in competitiveness in Sweden but not in Colombia, a country that scores substantially lower on gender equality, and Zhang (2011) finds no gender gap in the choice to compete in Chinese Han children but that girls are less competitive in both a matrilineal and a patrilineal group in China.

Contributing to these somewhat puzzling results, we find that in Armenia, a country that scores lower on macro gender equality indices than any of the above mentioned countries (see, for example, Hausmann et al., 2010), girls are in one task more competitive than boys when it comes to performance change, and in the other three tasks there are no gender differences.[23] There is no

[23]Global Gender Gap Report 2010 puts Armenia in rank 84, lower than China (61), Colombia (55), Israel (52), Austria (37) and Sweden (4).

gender difference in preferences for competition in any age group. As most previous studies, we find that girls are less risk taking than boys. Moreover, we find that girls are more altruistic than boys, something that is not always found among children. The risk taking results replicate what most other studies find on children and adolescents (and adults). Boys appear to become more risk taking than girls around the age of puberty in our sample, perhaps suggesting that the gender gap in risk taking is to an extent related to the hormones (Apicella et al., 2008; Sapienza et al., 2009), though see Zethraeus et al. (2009). Previous results on gender differences in altruism among children are mixed, thus more studies are needed on this topic.

Why girls are if anything more competitive and also less risk taking and more altruistic than boys in Armenia is for us somewhat of a puzzle. A natural extension of our study would be to perform a similar study in other countries of the former Soviet Union, in order to explore whether similar results would be observed. Another possible extension is to also study adults in a country such as Armenia, and also to study subjects living in the countryside rather in the capital, since gender related norms might differ between these areas. Moreover, it has previously been shown that the gender gap in competitiveness and risk taking partly depends on the institutional framework of the experiment (see Balafoutas and Sutter, 2010; Booth and Nolen, 2011a,b; Cason et al., 2010; Cotton et al., 2009; Ertac and Szentes, 2011; Niederle et al., 2010; Niederle and Yestrumskas, 2008; Wozniak et al., 2011). Testing these different setups in a large number of countries would provide an interesting venue for future research.

In sum, our results provide further evidence that it is important to perform studies in different samples, cultures and contexts in order to increase our understanding of (the development of) the gender gaps in economic preferences.

Appendix A.

In this appendix we present examples of the experimental tasks used in the competitiveness part of the second study (classroom part).

Table 12 shows a few examples of the math exercises that the children solved for various years. Students in years 2 to 5 had only to sum a random sequence of two two-digit numbers, while students in years 6 to 7 had to both add and subtract a random sequence of two two-digit numbers. Students in years 8 to 10 had to both add and subtract a random sequence of three two-digit numbers. All the numbers and mathematical operations were randomly generated to insure that the level of difficulty of the math task was the same throughout all the stages of the experiment for each of the year categories.

Table 12 – Examples of Math Tasks for Various Years

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

Figure 18 shows an example of the word search task that was used in the experiment. The students had to find and circle words in any direction on a straight line. Since these word search puzzles were not randomly generated by a computer there might have been slight differences in the difficulty of the word search task in different stages of the experiment. We used the same puzzle within each stage of the experiment for all the subjects.

Թ	Վ	Ի	Ց	Ե	Ղ	Ե	Կ	Ե	Ձ	Ճ	Ի	Յ
Ա	Ճ	Ը	Ի	Խ	Յ	Պ	Ո	Լ	Ո	Ռ	ՈՒ	Դ
Ն	Շ	ՈՒ	Դ	Գ	Օ	Թ	Ջ	Ք	Բ	ՈՒ	Չ	Յ
Գ	Ո	Ր	Գ	Ե	Ր	Լ	Մ	Փ	Ֆ	Ս	Ե	Օ
Ա	Կ	Ա	Ն	Ա	Ղ	Ե	Կ	Ի	Ր	Ե	Ս	Խ
Ր	Ե	Բ	Ա	Ր	Դ	Ի	Հ	Ծ	Ջ	Ր	Ո	Ն
Ա	Դ	Ա	Ս	Ա	Ր	Ա	Ն	Ի	Խ	Ե	Փ	Ա
Ն	Վ	Ֆ	Ղ	Ա	Յ	Ո	Ե	Շ	ՈՒ	Ն	Ի	Հ
Ա	Ն	Թ	Բ	Հ	Փ	Օ	Դ	Վ	Ք	ՈՒ	Մ	ՈՒ
Հ	Գ	Ի	Ն	Ե	Ր	Ո	Ձ	Ն	Խ	Ռ	Ե	Տ
Ա	Չ	ՈՒ	Հ	Բ	Թ	Ա	ՈՒ	Վ	Ա	Ա	Ճ	Ա
Ք	Ա	Ր	Ջ	Շ	Յ	Ջ	Կ	Ը	Ժ	Ղ	Ի	Պ
ՈՒ	Ճ	Ի	Ծ	Ն	Ա	Պ	Ա	Ս	Տ	Ա	Կ	Ֆ

Figure 18. Example of a word search task

Appendix B.

In this section we redo the analysis for the risk preferences employing the number of risky choices as a measure of risk taking. We replicate all the findings on risk taking in the main body of the paper. In the aggregate we find a significant gender gap in risk taking: boys take 20.89 percent more risk than girls do ($p < 0.001$). This gender gap, again, gets bigger with age. In the younger sample of school children (average age 8.83) the gap is 15.74 percent while in the older sample of school children (average age 14.72) it is 27.37 percent. The histograms in Figure 19 and information presented in Table 13 reveal that while in our sample of younger children the gender gap in mean risk taking is not significant, it is quite significant in the sample of older school children.

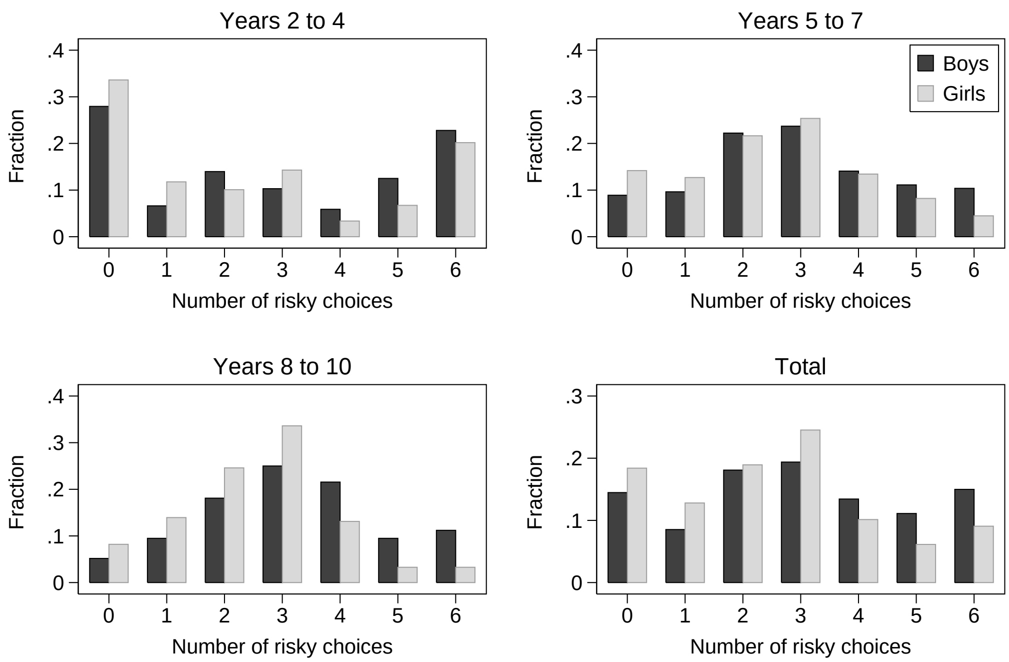

Figure 19. Distribution of risk as measured by the number of risky choices

Figure 20 presents the mean gender gaps in risk taking in an aggregated manner for all the three samples under study. Comparing older children (in years 8 to 10) with younger children (in years 2 to 4) we find that older children are more risk taking (Mann-Whitney U test: $p = 0.088$; Kolmogorov-Smirnov test: $p = 0.002$). While comparing older girls with younger girls and older boys with younger boys only the Kolmogorov-Smirnov test returns a significant test statistic ($p = 0.002$ and $p = 0.001$, respectively; Mann-Whitney $p-$values are 0.161 and 0.203, respectively).

Table 13 – Summary of Risk as Measured by the Number of Risky Choices

Age Group	Boys	Girls	All	M–W test (p–value)	K–S test (p–value)	t test (p–value)
Years 2 to 4	2.88 (0.20) 136	2.43 (0.21) 119	2.67 (0.15) 255	0.155	0.402	0.124
Years 5 to 7	2.99 (0.15) 135	2.54 (0.14) 134	2.77 (0.10) 269	0.040	0.535	0.027
Years 8 to 10	3.22 (0.15) 116	2.52 (0.12) 122	2.86 (0.10) 238	< 0.001	0.004	< 0.001
All	3.02 (0.10) 387	2.50 (0.09) 375	4.76 (0.07) 762	< 0.001	0.001	< 0.001

Notes: The table shows the mean as the main number and the standard deviation in parenthesis. Number of observations are below the means.

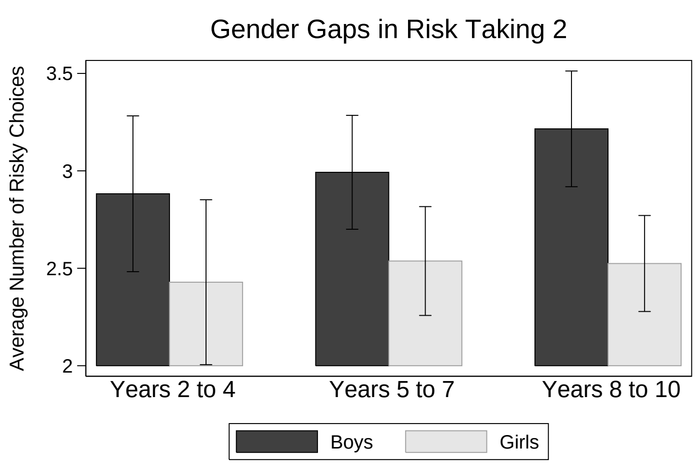

Figure 20. Average number of risky choices, by gender and age group

References

Andersen, S., S. Ertac, U. Gneezy, J. A. List, and S. Maximiano (2011): "Gender, Competitiveness and Socialization at a Young Age: Evidence from a Matrilineal and a Patriarchal Society," Unpublished.

Apicella, C. L., A. Dreber, B. C. Campbell, P. Gray, M. Hoffman, and A. C. Little (2008): "Testosterone and Financial Risk Preferences," *Evolution and Human Behavior*, 29, 384–390.

Balafoutas, K. and M. Sutter (2010): "Gender, Competition and the Efficiency of Policy Interventions," IZA Discussion Paper 4955.

Benenson, J. F., J. Pascoe, and N. Radmore (2007): "Children's Altruistic Behavior in the Dictator Game," *Evolution and Human Behavior*, 28, 168–175.

Bertrand, M. (2010): "New Perspectives on Gender," in *Handbook of Labour Economics*, Elsevier, vol. 4b, chap. 17, 1545–1592.

Blake, P. R. and D. G. Rand (2010): "Currency Value Moderates Equity Preference Among Young Children," *Evolution and Human Behavior*, 31, 210–218.

Booth, A. L. and P. J. Nolen (2011a): "Choosing to Compete: How Different are Girls and Boys?" *Journal of Economic Behavior and Organization*, in press.

——— (2011b): "Gender Differences in Risk Behaviour: Does Nurture Matter?" *Economic Journal*, in press.

Borghans, L., B. Golsteyn, J. J. Heckman, and H. Meijers (2009): "Gender Differences in Risk Aversion and Ambiguity Aversion," *Journal of the European Economic Association*, 7, 649–658.

Cardenas, J.-C., A. Dreber, E. von Essen, and E. Ranehill (forthcoming): "Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden," *Journal of Economic Behavior and Organization.*

Cason, T. N., W. A. Masters, and R. M. Sheremeta (2010): "Entry into Winner-Take-All and Proportional-Prize Contests: An Experimental Study," *Journal of Public Economics*, 94, 604–611.

Cotton, C., F. McIntyre, and J. Price (2009): "Gender Differences Disappear with Exposure to Competition," Unpublished.

Croson, R. and U. Gneezy (2009): "Gender Differences in Preferences," *Journal of Economic Literature*, 47, 448–54.

Dreber, A., E. von Essen, and E. Ranehill (2011a): "In Bloom: Gender Differences in Preferences among Adolescents," Unpublished.

——— (2011b): "Outrunning the Gender Gap – Boys and Girls Compete Equally," *Experimental Economics*, 14, 567–582.

Eckel, C. C. and P. J. Grossman (2008a): "Differences in the Economic Decisions of Men and Women: Experimental Evidence," in *Handbook of Experimental Economics Results*, ed. by C. R. Plott and V. L. Smith, Elsevier, vol. 1, chap. 57, 509–519.

——— (2008b): "Men, Women and Risk Aversion: Experimental Evidence," in *Handbook of Experimental Economics Results*, ed. by C. R. Plott and V. L. Smith, Elsevier, vol. 1, chap. 113, 1061–1073.

Engel, C. (2011): "Dictator Games: A Meta Study," *Experimental Economics*, 14, 583–610.

Ertac, S. and B. Szentes (2011): "The Effect of Information on Gender Differences in Competitiveness: Experimental Evidence," Unpublished.

Fehr, E., D. Rutzler, and M. Sutter (2011): "The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence," IZA Discussion Paper 5530.

Gneezy, U., K. L. Leonard, and J. A. List (2009): "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society," *Econometrica*, 77, 1637–1664.

Gneezy, U., M. Niederle, and A. Rustichini (2003): "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, 118, 1049–1074.

Gneezy, U. and A. Rustichini (2004a): "Executives verses Teachers: Gender, Competition and Self Selection," Unpublished.

——— (2004b): "Gender and Competition at a Young Age," *American Economic Review: Papers and Proceedings*, 94, 377–381.

Gummerum, M., Y. Hanoch, M. Keller, K. Parsons, and A. Hummel (2010): "Preschoolers' Allocations in the Dictator Game: The Role of Moral Emotions," *Journal of Economic Psychology*, 31, 25–34.

Gunther, C., N. A. Ekinici, C. Schwieren, and M. Strobel (2009): "Women Can't Jump? An Experiment on Competitive Attitudes and Stereotype Threat," *Journal of Economic Behavior and Organization*, 75, 395–401.

Gupta, N. D., A. Poulsen, and M.-C. Villeval (forthcoming): "Male and Female Competitive Behavior – Experimental Evidence," *Economic Inquiry*.

Harbaugh, W. T., K. Krause, and S. Liday (2003): "Bargaining by Children," Unpublished.

Harbaugh, W. T., K. Krause, and L. Vesterlund (2002): "Risk Attitudes of Children and Adults: Choices Over Small and Large Probability Gains and Losses," *Experimental Economics*, 5, 53–84.

Hausmann, R., L. D. Tyson, and S. Zahidi (2010): *The Global Gender Gap Report 2010*, Geneva, Switzerland: World Economic Forum.

Kamas, L. and A. Preston (2010): "Social Preferences, Competitiveness and Compensation: Are there Gender Differences?" Unpublished.

Niederle, M., C. Segal, and L. Vesterlund (2010): "How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness," Unpublished.

Niederle, M. and L. Vesterlund (2007): "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, 122, 1067–1101.

Niederle, M. and A. H. Yestrumskas (2008): "Gender Differences in Seeking Challenges: The Role of Institutions," Unpublished.

Niels and G. Reiner (2010): "Explaining Gender Differences in Competitiveness: Gender-Task Stereotypes," Jena Economic Research Papers 2010 - 017.

Sapienza, P., L. Zingales, and D. Maestripieri (2009): "Gender Differences in Financial Risk Aversion and Career Choices are Affected by Testosterone," *Proceedings of the National Academy of Sciences*, 106, 15268–15273.

Shurchkov, O. (forthcoming): "Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints," *Journal of the European Economic Association*.

Sutter, M., M. G. Kocher, D. Rutzler, and S. T. Trautmann (2010): "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior," IZA Discussion Paper 5404.

Sutter, M. and D. Rutzler (2010): "Gender Differences in Competition Emerge Early in Life," IZA Discussion Paper 5015.

Wozniak, D., W. T. Harbaugh, and U. Mayr (2011): "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices," MPRA Paper No. 31374.

Zethraeus, N., L. Kocoska-Maras, T. Ellingsen, B. von Schoultz, A. L. Hirschberg, and M. Johannesson (2009): "A Randomized Trial of the Effect of Estrogen and Testosterone on Economic Behavior," *Proceedings of the National Academy of Sciences*, 106, 6535–6538.

Zhang, J. (2011): "Do Girls in China Compete Just As Much As Boys? Evidence from an Experiment that Predicts Educational Choice," Unpublished.