Dreber, Anna; von Essen, Emma; Ranehill, Eva

Working Paper

Outrunning the gender gap: Boys and girls compete equally

SSE/EFI Working Paper Series in Economics and Finance, No. 709

Provided in Cooperation with:
EFI - The Economic Research Institute, Stockholm School of Economics

Outrunning the Gender Gap –

Boys and Girls Compete Equally*

Anna Dreber[†‡], Emma von Essen[§], and Eva Ranehill[♪]

SSE/EFI Working Paper Series in Economics and Finance
No 709
March 2011

Abstract

Recent studies find that women are less competitive than men. This gender difference in competitiveness has been suggested as one possible explanation for why men occupy the majority of top positions in many sectors. In this study we explore competitiveness in children, with the premise that both culture and gendered stereotypes regarding the task at hand may influence competitive behavior. A related field experiment on Israeli children shows that only boys react to competition by running faster when competing in a race. We here test if there is a gender gap in running among 7-10 year old Swedish children. We also introduce two female sports, skipping rope and dancing, to see if competitiveness is task dependent. We find no gender difference in reaction to competition in any task; boys and girls compete equally. Studies in different environments with different types of tasks are thus important in order to make generalizable claims about gender differences in competitiveness.

Keywords: competitiveness; gender differences; field experiment.

JEL codes: C93, D03, J16.

*We are very grateful for comments from Johan Almenberg, Anne Boschini, Hannah Riley Bowles, Armin Falk, Uri Gneezy, Moshe Hoffman, Magnus Johannesson, Astri Muren, Paul Nystedt, Mats Persson, David G. Rand and seminar participants at the Gender in Negotiation and Decision Making Seminar at Harvard Kennedy School, Harvard Business School, Harvard University, Ratio Institute, Santa Fe Institute, Stockholm School of Economics, Stockholm University, the 3rd Nordic Conference on Behavioral and Experimental Economics, University of Bonn and University of Zürich. Financial support from the Jan Wallander and Tom Hedelius Foundation (A.D., E.R.) and the Carl Silfvén Foundation (E.R) is gratefully acknowledged, as well as financial support from the Swedish Council for Working Life and Social Research (FAS).
[†]Institute for Financial Research (SIFR), Stockholm. Email: anna.dreber@sifr.org
[‡]Women and Public Policy Program, Harvard Kennedy School
[§]Department of Economics, Stockholm University. Email: emma.vonessen@ne.su.se
[♪]Department of Economics, Stockholm School of Economics. Email: eva.ranehill@hhs.se

1. Introduction

Men occupy the majority of top positions in most societies, both in the private and in the public sector. The proposed reasons for this remain highly controversial within academia as well as politics (Ceci and Williams 2006). Today, women in many countries are at least as likely as men to pursue higher education, and female labor force participation has risen to levels similar to that of men. Meanwhile, a number of recent studies show that women compete less than men. Competitiveness is typically measured as either a preference for competition, such as self-selecting into a tournament instead of a piece-rate payment scheme, or by the performance response as a reaction to a competitive setting compared to a non-competitive setting. Many studies find that only males perform better under competition (Gneezy et al. 2003, Gneezy and Rustichini 2004a), or that when both men and women perform better, males still perform significantly better than women (Datta Gupta et al. 2005). It has also been shown that women tend to prefer the non-competitive setting even when there is no gender gap in performance in the competitive setting and that men compete more than what is optimal for them while women compete less (Niederle and Vesterlund 2007). Some studies find that competitiveness depends on the gender of the opponent(s) (Gneezy et al. 2003, Gneezy and Rustichini 2004a, Datta Gupta et al. 2005, Price 2008) whereas some find that women's competitiveness depend on the institutional framework (e.g., Niederle and Yestrumskas 2008, Niederle et al. 2009, Balafoutas and Sutter 2010). These gender differences have been suggested as a possible explanation for the gender gap in the labor market. The policy implications of a gender gap in competitiveness depend on the causes of the gap. Whether these gender differences are innate or acquired later in life remains unknown. Children therefore provide an interesting subject pool for the study of this distinction.

In this paper, we explore whether there are gender differences in competitiveness among children. Two previous studies also investigate this. Booth and Nolen (2009) look at willingness to compete in solving mazes among adolescent boys and girls from single sex schools and from mixed schools. Boys compete equally in both schools and more than girls do, whereas girls in single sex schools compete more than girls from mixed schools. In a field experiment looking at 9-10 year old Israeli children, Gneezy and Rustichini (2004a) find that boys, but not girls, respond to competition by running faster against another child than when

running alone. Moreover, they find that the gender of the opponent matters only for girls, who compete less when running against another girl.

In this study we run a field experiment on 7-10 year old children in Sweden. The design is inspired by that of Gneezy and Rustichini (2004a), where the children compete in running. In addition, in our study the children also compete in skipping rope and dancing. The running task is included in order to have a comparison to previous work, while varying country (Israel vs Sweden), even though some parameters differ between the two studies. The other two tasks are included to study whether there are male and female areas of competition. If tasks are gendered, it is possible that this leads to gender differences in both motivation for, and payoffs from, competing. Most competitiveness studies build on tasks such as solving mazes and performing simple arithmetic, which are generally considered as male tasks. Several studies show that women perform worse on standardized tests when they are reminded of negative stereotypes about female math ability (Inzlicht and Ben-Zeev 2003, O'Brien and Crandall 2003, Shih et al. 1999, Steele 1997).[1] This kind of stereotype threat has been suggested as one reason why women in mixed gender groups compete less than men in some of the tasks previously studied in this literature (Gneezy et al. 2003). There is mixed evidence on the role of the task on the gender gap in competitiveness. Günther et al. (2009) and Grosse and Reiner (2010) find a gender difference in performance change in a math task but not in a word task, whereas Wozniak et al (2010), using a maze task and the same word task, find no difference in the gender gap between the two tasks.[2] Thus, to explore competitiveness more generally than what has previously been done, in particular on children, we also look at what we consider more female tasks. Since our experiment is conducted with children, our inspiration comes from tasks that children perform. The tasks were chosen in agreement with the teachers.

Competitiveness is measured in the same way for all three tasks. First the children perform the task individually. Their performance is measured and they are then matched together in pairs of two depending on their result. Thereafter the children perform the task a second time in these matched pairs. Competitiveness is measured as the difference in performance between the individual and matched performance, and is thus considered as the reaction to

[1] Interestingly when women are told that there are no differences between men and women in abstract math tests, women perform as well as men (Spencer et al. 1999).

[2] Wozniak et al. (2010) find no gender gap in performance change but find that men are more likely to self-select into competitions.

competition. We have a control group of children who perform the task alone a second time, as in Gneezy and Rustichini (2004a). This allows us to control for unobservable factors that could cause differences in the outcome, such as, e.g., one gender getting tired faster than the other.

Given previous literature, we hypothesize that if there is a gender gap in running, boys will compete more than girls. We also hypothesize that if there is a gender gap in the female tasks it will be the opposite since, if anything, these tasks have positive stereotypes regarding female ability.

We find no evidence in support of our hypotheses. We find no gender differences in competitiveness among children in Sweden in any of the three tasks. Boys and girls increase their performance equally in the competitive setting for running and skipping rope, and there is no difference between the average increases. Regarding the dancing task, both boys and girls decrease their performance when competing, and this decrease in performance is not significantly different between the two genders. Our results also indicate that the gender of the opponent does not alter performance of either gender in any of the three tasks. Moreover, the findings from the control group indicate that our results are not driven by gender differences in factors such as tiredness.

This contradiction to earlier results by Gneezy and Rustichini (2004a) may be explained by context, such as culture. It has previously been shown that cultural factors such as gender norms may influence competitive behavior. Gneezy et al. (2009) compare a matrilineal society in India with a patriarchal society in Tanzania and find that women prefer the competitive setting more than men in the matrilineal society, whereas the inverse is found in the patriarchal society.[3] Our results suggest that cultural factors matter also among Western countries. Even though we cannot directly test this, we speculate that the difference between our results and those of Gneezy and Rustichini (2004a) may be due to differences in gender norms. Even though Sweden and Israel are both Western societies with high female labor force participation, Sweden usually performs higher on gender equality indices.[4]

[3] The task at hand is the toss of a tennis ball into a bucket. Gneezy et al. (2009) are unaware of any resemblance between this task and some popular task in the cultures that are being studied, thus it is unlikely that the specific task had a certain gendered stereotype. In general, however, throwing objects could be considered more male in many cultures since men have typically been the hunters (e.g., men hunt through spear throwing).

[4] The Global Gender Gap Report 2009 lists Sweden as number four in the world in terms of gender equality. Israel ranks 45th out of 134 countries.

Our paper is organized as follows. In section 2, we describe the experimental design of our field study. In section 3, we present our results. We conclude in section 4, where we also discuss the possible explanations for our findings as well as promising directions for future research.

2. Experimental design

The field experiment was conducted in 11 primary school classes in the Stockholm area during 2008 and 2009. We contacted all primary schools in Stockholm with a letter explaining that we intended to study competitiveness among children. There was no mentioning of the gender dimension. All tasks were performed during physical education classes and the experiment was overseen by the teacher. The children, aged 7-10 years old, did not realize that they were participating in an experiment (as in Gneezy and Rustichini 2004a). The teachers did not mention the study to the children, and the tasks are standard in Swedish physical education classes. On two or three different occasions, the children competed in running, skipping rope and modern dance. These three tasks were carefully chosen. Running has previously been explored in Gneezy and Rustichini (2004a) and is part of physical education in Sweden. Skipping rope is a task that girls perform during school breaks throughout the world, including Sweden. Dancing is often considered female (Henschel-Pellet 2001), and during the Swedish school year it typically takes up one physical education class. The running task was administered by the teachers on a separate occasion (as in Gneezy and Rustichini 2004a), whereas the skipping rope was instructed and administered by the experimenters as an exercise complementary to the dancing, thus the experimenters were present at the occasion the dance competition. The dancing task was designed, instructed and scored by a professional dance teacher on one or two occasions depending on the length of the class. To avoid that teachers treated boys and girls differently in order to affect the results of the study, all teachers, including the dance teacher, were unaware of the gender dimension of the study. The children were given 40 minutes to practice the dancing task together with the whole class and the dance teacher, and 5 minutes to practice the skipping rope task prior to the start of the experiment.

In running, performance is measured by how fast the children ran 60 meters, the distance normally used for short distance running in Swedish schools. Note that this distance differs from what Gneezy and Rustichini (2004a) used, 40 meters. In skipping rope (where two individuals turn the rope while one child jumps), performance was measured as the number

of jumps performed until the children missed. In dancing, the dance teacher scored the children based on how they performed compared to the set goal of the dance choreography. The dance choreography included ten distinct exercises and the children were awarded one point for each of these ten movements that they performed correctly.[5]

Each task consisted of two stages. At the first stage, the children performed the task by themselves and individual performance was measured. The teachers were aware of the setup of the study, whereas the children were unaware of the existence of a second stage when performing the task in the first stage in all three tasks.[6] At the second stage, the children performed the task in competition with another child. Matching started with the two children that had the best performance in the first stage in each task, and then continued down the list. If more than two children obtained the same result in the first stage, matching was done randomly (as in Gneezy and Rustichini 2004a).[7] In the case of dancing, both the individual performance and the competition occurred in a separate room where only the one or two children dancing and dance teacher were present. In all three tasks, the children knew that their competitor had achieved a similar score at the first stage. The dance teacher presented the tasks as competitive activities. The dance competition was presented as a "battle", somewhat in the spirit of a popular TV show.[8] In the skipping rope task, two ropes were put next to each other. The children were instructed to start jumping at the same time and were told that the winner was the child who performed the greatest number of jumps. All rules were explained by the dance teacher and the experimenters and no compensation was awarded apart from the intrinsic motivation that comes from winning, as in Gneezy and Rustichini (2004a). Our measure of competitiveness is the change in performance between the first and the second stage of the tasks.

[5] The dancing task consisted of a one minute long modern dance phrase. The choreography of the phrase was focusing on strength, coordination and balance rather than "feminine grace", in order to minimize subjectivity in the evaluation of dance. Since the dance teacher was not aware of the purpose of the study, any potential subjectivity is likely to be orthogonal to the gender of the child evaluated. The children were aware of how the task was scored.

[6] The teachers were aware of the two stages of each task, but did not inform the children about this. The experimenters gave oral instructions to the children about the setup of the study at the relevant stages.

[7] When an unequal number of children performed equally well, they were randomly paired. The remaining child was matched with the child with the next best result. If more than one child had the next best result, the remaining child with the higher score from the first matching was randomly matched with one of these children. During the competitive part of the experiment, the competing pairs participated in random order.

[8] The TV show "So you think you can dance" was aired on Swedish television before and during the time the study was performed. It has been pointed out to us that dancing is often a cooperative or communal activity. We assume that the competitive element of the TV show decreased the cooperative or communal aspects of the dancing task.

3. Results

We test whether there is a gender gap in competitiveness among children in Sweden and whether the nature of the task affects the size and direction of the gender gap. We start by looking at gender differences in competitive behavior. Thereafter we address the effect of the gender composition in the competitive setting. We also present a robustness check and a survey on how boyish/girlish children perceive the explored tasks to be. For all tests in the analysis, we have performed a Mann-Whitney test, a two-sided t-test and used bootstrap techniques. Throughout the analysis we present only the p-value for the Mann-Whitney test.[9]

3.1 No significant gender differences in competitive behavior

In our study, 149 children participated in running, 143 in skipping rope, and 146 in dancing. The gender distribution in the three sports was 68 boys and 81 girls in running, 67 boys and 76 girls in skipping rope and 64 boys and 82 girls in dancing.[10] Consistent with sex-stereotypic expectations, we find that in the individual setting (stage 1) boys ran on average faster than girls (unlike in Gneezy and Rustichini 2004a), and girls skipped rope better compared to boys. In running and skipping rope, the p-value for a significant gender difference is 0.008, with boys performing better in the former ask and girls in the latter. In dancing, the non parametric test gives a p-value of 0.0478, whereas the difference is not significant with a t-test or a bootstratpped test.[11] When it comes to competitiveness, table 1 below shows that in all three tasks, and for both genders, average performance in the competitive setting differs significantly from average performance in the non-competitive

[9] We present the Mann-Whitney test since none of our variables are normally distributed when using a skewness and kurtosis test. When there is a difference between the parametric and non parametric tests in terms of significance we also report the p-values for the t-test and the bootstrap-based critical values. We have also compared whether the distributions for each reported variable differ between men and women using a Kolmogorov-Smirnov test. The results are the same as those reported for mean values.

[10] Two subjects, one boy and one girl, were dropped from the sample due to physical disabilities. The differences in number of children between activities are due to the fact that we had different number of occasions depending on the structure of the physical education classes in the different schools. There is no significant difference in performance change between school classes that had one occasion or school classes that had more occasions (ranksum: p=0.53).

[11] When we perform the tests on the inner quartile range (IQR, the distribution between the 25^{th} and 75^{th} percentile) the Mann-Whitney test is also insignificant.

setting, ($p<0.01$). Both genders improve their performance significantly in running and skipping rope in the competitive setting, but perform worse in dancing.[12]

	Running		SR	Skipping rope		SR	Dancing		SR
	Stage 1	Stage 2	p-value	Stage 1	Stage 2	p-value	Stage 1	Stage 2	p-value
Girls	11.92	11.66	0.000	49.01	69.37	0.000	5.87	5.13	0.001
Boys	11.55	11.42	0.002	32.48	45.12	0.000	5.27	4.48	0.001

Table 1. Average performance in stage 1 and in stage 2. Signrank (SR) test p-values of performance change for girls and boys separately.

Figures 1-3 below show the distribution of the performance change in the different tasks. The three histograms show that there are no significant gender differences in any of the three tasks (running: $p=0.47$, skipping rope: $p=0.24$, dancing: $p=0.85$).[13]

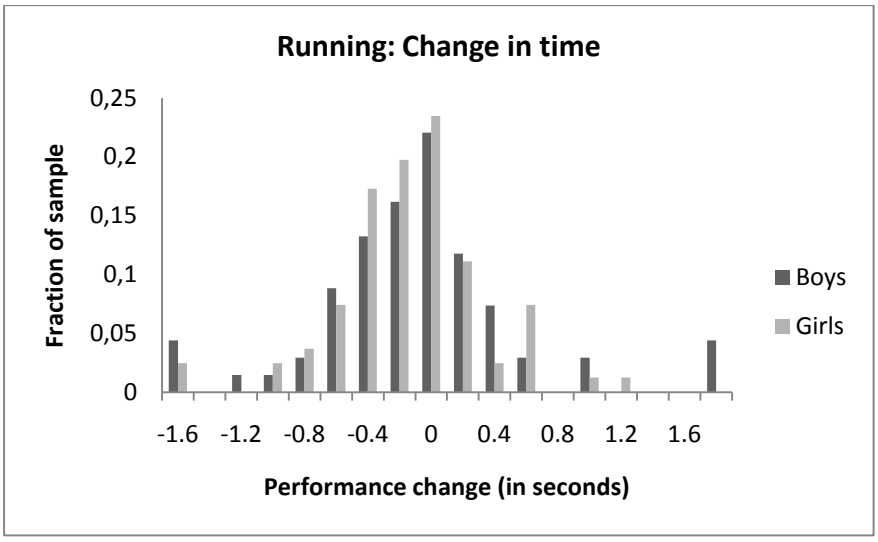

Figure 1. Distribution of change in running time (stage 2 - stage 1), by gender.

[12] The other tests are not significant when it comes to the performance change of boys in running. However, when performing the tests on the IQR, all three tests are significant.

[13] To further investigate a possible gender difference in performance change we also performed quantile regressions for each task, controlling for gender of opponent (performed for quantile 0.1-1.0). Gender has an effect only in the top 10% of the performance change distribution in running and skipping rope. In this part of the distribution the performance change of boys is larger than girls in running and the opposite for skipping rope. There are however very few observations in the top 10% for each task. These results are therefore mere indications.

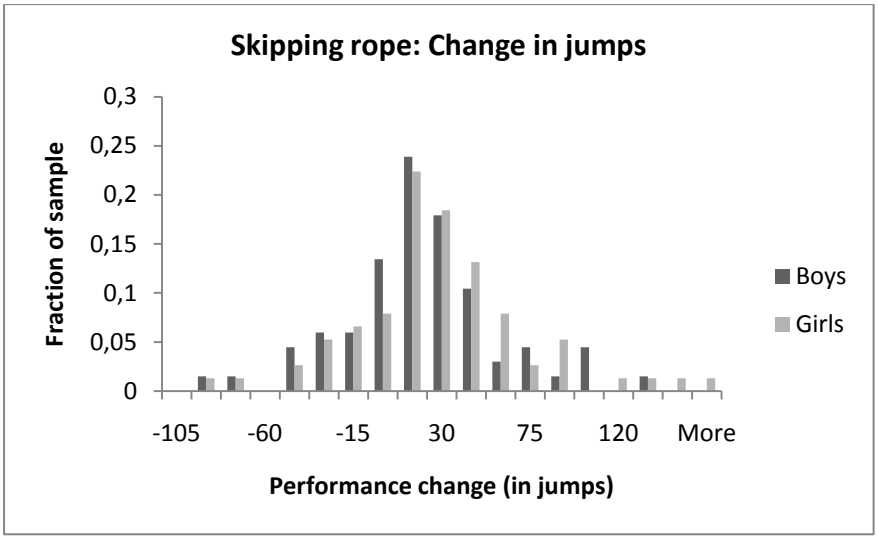

Figure 2. Distribution of change in jumps (stage 2 – stage 1), by gender.

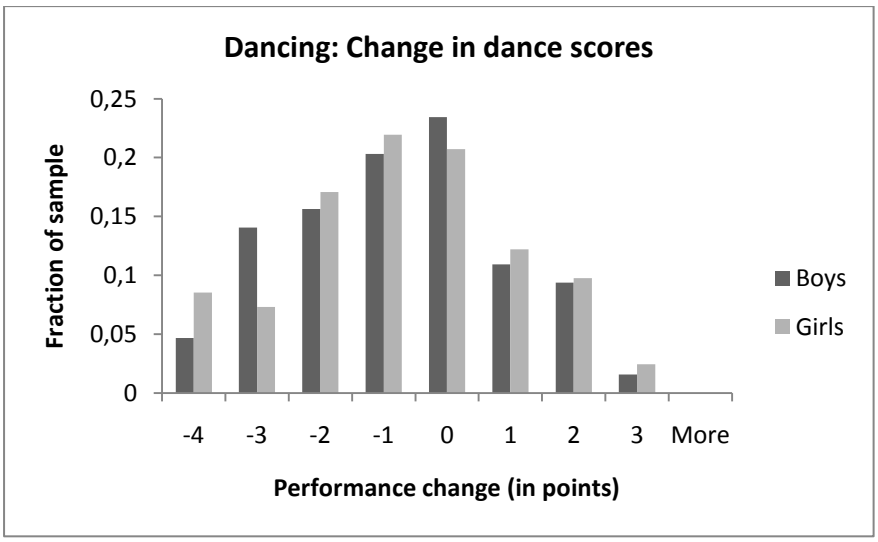

Figure 3. Distribution of change in dance scores (stage 2 – stage 1), by gender.

The pattern of gender similarities are displayed in an aggregated manner in figures 4-6 below. These plots show the average change in performance by each gender. In running, girls improve on average 0.26 seconds, or about 2.1%. This can be compared to the average decrease in running time of 0.13 seconds, or 1.1%, for boys.[14] The corresponding numbers for skipping rope is an increase of 20 versus 13 jumps, implying an improvement of 42% and 39% respectively. On average, girls' dance performance deteriorates by 0.73 points (13%) on average and boys' by 0.78 points (15%). As stated above, the difference in average change in

[14] For all three tasks, we conducted the same analysis with relative performance, where relative performance was defined as ((stage2-stage1)/stage1). This did not change any of our results. Our findings further remain stable when excluding outliers. An outlier is defined as an observation that lie more than two standard deviations away from the sample mean.

performance between boys and girls is not statistically significant in any of the three cases.[15] These results also hold within all age groups in our sample.[16]

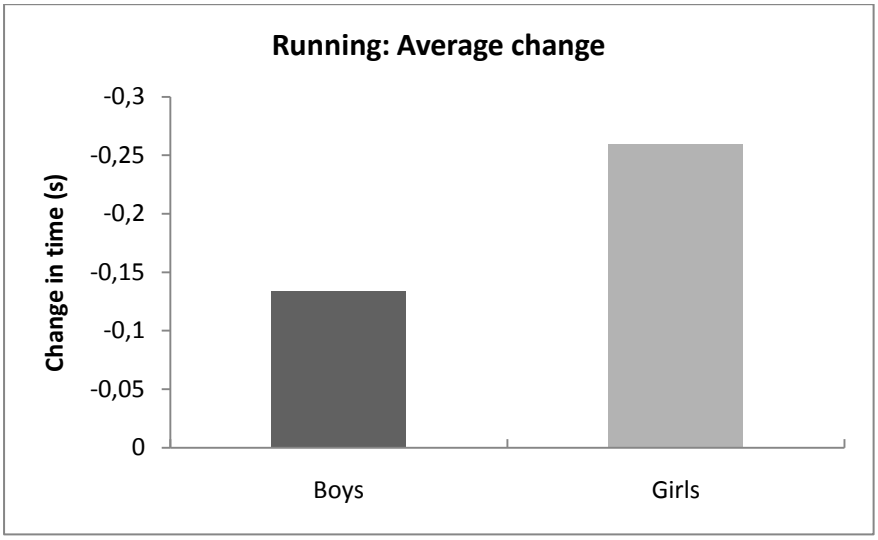

Figure 4. Average change in time (stage 2 – stage 1), by gender. 78 girls and 71 boys.

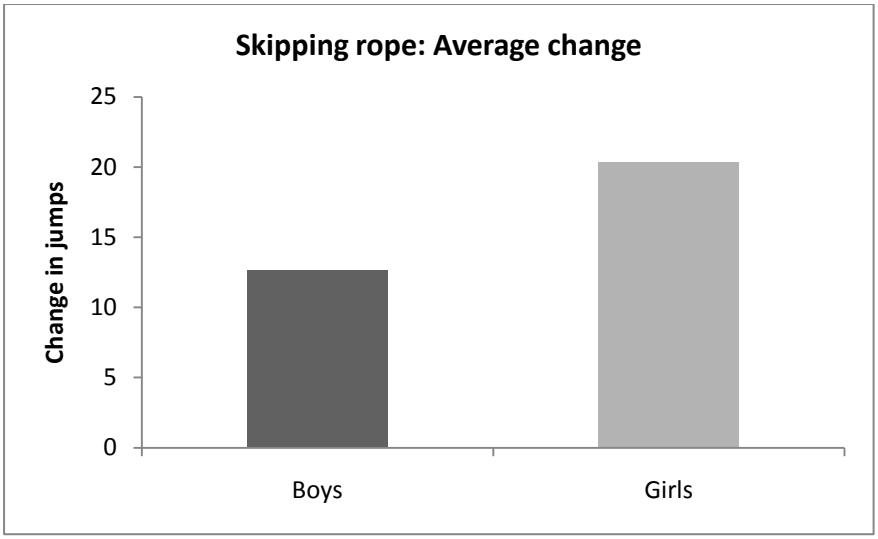

Figure 5. Average change in jumps (stage 2 – stage 1), by gender. 74 girls and 69 boys.

[15] A sample size analysis indicates that 1411, 965 and 38407 observations would be needed to obtain a significant result for the performance change in running, jumping and dancing respectively. The basis for the power calculation is a significance level of 5% and a power of 80%.

[16] In particular, when we restrict the running analysis to the same age group as studied in Gneezy and Rustichini (2004a), the gender gap among these 114 children aged 9-10 years old is still insignificant (p=0.47).

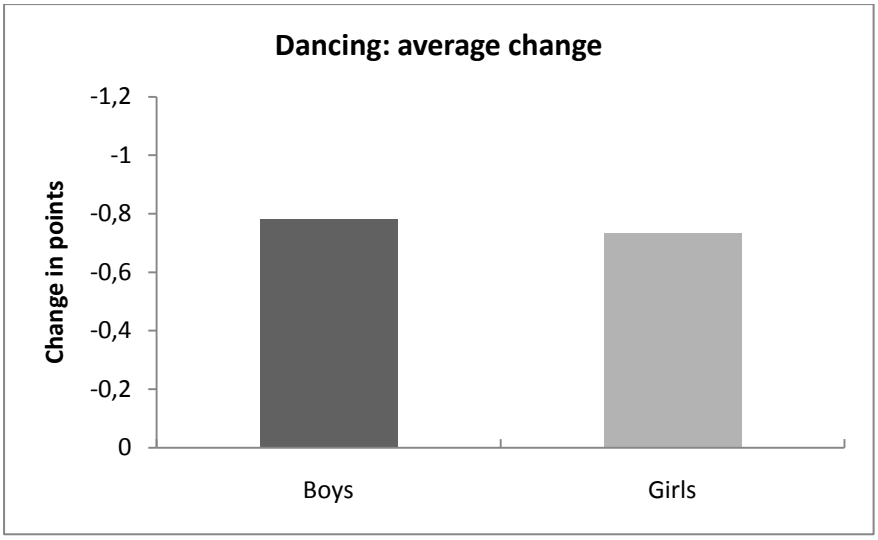

Figure 6. Average change in dance scores (stage 2 – stage 1), by gender. 82 girls and 64 boys.

3.2 Impact of opponent gender on competitive behavior

Some previous studies find that women compete more against women, and men more against men (e.g., Datta Gupta et al. 2005, Gneezy et al. 2003). On the contrary, Gneezy and Rustichini (2004a) find that boys are not affected by the gender composition but girls compete more against boys. Our results suggest that neither boys nor girls are influenced by the gender of their opponent. Table 2 gives an overall summary of our results for the different pair compositions in our study. In running, both girls and boys improve the most when running against a girl. However, the difference in competitive behavior when facing the same vs facing the opposite gender is statistically insignificant for girls ($p=0.6221$) and for boys ($p=0.0701$). In skipping rope and dancing, girls compete more fiercely against boys, but none of these results are significant (skipping rope: $p=0.1864$, dancing: $p=0.4982$). Boys on the other hand compete more against boys in skipping rope and more against girls in dancing, though also these differences are not significant (skipping rope: $p=0.8401$, dancing: $p=0.4519$).

	Running			Skipping rope			Dancing		
Sample	**N**	**Stage2-stage1**	**p-value**	**n**	**Stage2-stage1**	**p-value**	**n**	**Stage2-stage1**	**p-value**
Total	149	-0.20	0.000	143	17	0.000	146	-0.75	0.000
Girls with girls	47	-0.28	0.001	40	14	0.026	41	-0.83	0.002
Boys with boys	42	-0.13	0.175	30	15	0.014	27	-0.96	0.005
Girls mixed pairs	34	-0.24	0.001	36	27	0.001	41	-0.63	0.079
Boys mixed pairs	26	-0.14	0.001	37	10	0.127	37	-0.65	0.054

Table 2. Performance change (stage 2 – stage 1) based on the gender composition of the competing pairs.

3.3 Robustness checks

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 getting tired faster than the other. The control group includes 66 children in the running task (31 boys and 35 girls), 65 children in the skipping rope task (29 boys and 36 girls), and 49 children in the dancing task (19 boys and 30 girls). For running, both boys and girls perform worse in stage 2 compared to stage 1 ($p<0.001$). Importantly, however, there is no significant gender difference when we test performance change between boys and girls ($p=0.4878$). The fact that stage 2 performance in running is worse than stage 1 performance indicates an even greater reaction to competition in running for both boys and girls than if there would have been no performance change in the control. The absolute performance change between stage 2 and stage 1 in skipping rope and dancing is not significant (skipping rope: $p=0.1627$, dancing: $p=0.3206$). This indicates that when not competing against another child there is no significant improvement in performance in these two tasks. Moreover, there are no significant differences in these two tasks when we test performance change between boys and girls (skipping rope: $p=0.9106$, dancing: $p=0.9664$). See table 3 for more details on the results.

Control		Running			Skipping rope			Dancing	
Sample	**n**	**Stage2-stage1**	**p-value**	**N**	**Stage2-stage1**	**p-value**	**n**	**Stage2-stage1**	**p-value**
Total	66	0.35	0.001	65	6.77	0.163	49	-0.35	0.321
Gender difference	66	-0.20	0.488	65	-3.69	0.911	49	0.22	0.966

Table 3. Performance change (stage 2 – stage 1) in the control, and whether there is a gender difference in this performance change.

Even though we find no significant gender differences in mean change in performance in our main analysis, there may be differences in the variances of the performance distributions. We test this and find no significant differences in the variance of change in performance between boys and girls.[17]

Furthermore, we also perform a within subject analysis across tasks. We balance the sample by keeping only individuals that performed all three tasks (58 girls and 45 boys). We find no correlations between performance change in the different tasks for boys or girls (running and skipping rope: boys: p=0.5058, girls: p=0.3617; running and dancing: boys: p=0.4389, girls: p=0.9088; skipping rope and dancing: boys p=0.2710, girls: p=0.1089).[18] This suggests that in our sample there does not seem to be a general competitive type – some individuals perform better under competition in one task and not another.

3.4 Do children perceive the tasks to be gendered?

In a separate survey of children aged 9-10 years old, we asked how boyish/girlish they considered running, skipping rope and dancing to be. We also elicited perceptions of how boyish/girlish competing in these tasks was. The children were asked to use a scale where a lower number indicates rating the task as more boyish and a higher number as more girlish (1=very boyish, 2=boyish, 3=neutral, 4=girlish, 5=very girlish).

Table 4 shows that, on average, running is perceived to be more boyish than skipping rope and dancing. This is the case both in absolute and relative terms.

[17] The most common test for comparison of standard deviations, the F-test for the homogeneity of variances (sdtest), is very sensitive to the assumption that that the data are drawn from an underlying normal distribution. Therefore we also performed a robust test (Levene's test with mean, median and 10% trimmed mean). None of these tests indicated significant differences in the variances.

[18] Performing this analysis on relative performance change does not alter the results qualitatively.

Variable	Obs	Mean	Std Dev	Min	Max
Running	34	2.68	0.73	1	4
Skipping rope	35	4.17	0.79	3	5
Dancing	34	4.03	0.83	2	5
Competition running	35	2.29	0.83	1	4
Competition skipping rope	35	3.77	0.94	2	5
Competition dancing	35	4.03	0.82	3	5

Table 4. Summary statistics of ratings.

Running is perceived as significantly more boyish than skipping rope ($p<0.001$) and dancing ($p<0.001$).[19] When comparing skipping rope and dancing there is no significant difference ($p=0.5432$). When it comes to the perceptions of how boyish/girlish it is to compete in these tasks, we observe the same pattern. Competing in running is rated as more boyish than competing in skipping rope and dancing

We also compare the rating of competing in a certain task with the general rating of the task. Competition in itself is rated as more boyish compared to the general rating for both running and skipping rope ($p=0.0315$ and $p=0.0211$), but not for dancing. For dancing there is no significant difference between competition and the general rating of the task ($p=1$). When merging these data, competition seems to be rated more boyish compared to the rating of the task in general ($p=0.0050$).

3.4.1. Do boys and girls have different perceptions?

In table 5 we divide the ratings by gender. Girls tend to rate running as gender neutral and boys as more boyish ($p=0.0021$). Moreover, girls tend to rate dancing as more neutral, whereas boys rate it as more girlish ($p=0.0430$). Boys and girls give skipping rope a similar score. Regarding competition, there is no significant difference in the ratings for any of the tasks.

	Running	Skipping rope	Dancing	Competition running	Competition skipping rope	Competition dancing
Girls	3.06	4	3.81	2.53	3.65	3.88
Boys	2.31	4.35	4.35	2.06	3.82	4.18
Total	2.70	4.18	4.09	2.29	3.74	4.03

Table 5. Average ratings by gender.

[19] Most of these variables are not normally distributed according to a skewness and kurtosis test. Thus, we perform a Mann-Whitney test for differences in distributions between the tasks.

When merging the data on the three tasks, girls and boys rate competition in the same way in terms of how boyish/girlish it is (p=0.6993).[20]

4. Discussion

Previous literature on competitive behavior finds that men compete to a larger extent than women. This difference in behavior may explain part of the gender gap observed in many areas in society. In this literature, however, only a few tasks have been used to measure competitiveness, and these tasks can arguably be considered as more male than female. Three studies find that gender differences in competitiveness vary with the task at hand (Gneezy and Rustichini 2004b, Günther et al. 2009, Grosse and Reiner 2010), whereas another study find no difference in the gender gap between a maze task and a word task (Wozniak et al. 2010). Meanwhile, work in social psychology suggests that individual perceptions about relative performance, such as (over)confidence, and especially stereotypes may have important implications for actual performance (Steele 1997, Shih et al. 1999). Exploring more tasks than maze solving and simple arithmetic is thus important in order to increase our understanding about gender differences in competitiveness and the potential role of stereotypes.

In this paper we study how children compete in three distinct tasks. We let the children compete in running in order to create a comparison with previous literature. Moreover, we add two more female tasks to the competition; skipping rope and dancing. Competitiveness is measured by reaction to competition, i.e. as the child's increase in performance when competing against another child, compared to when the task is performed individually. We find no gender differences in competitive behavior in any of these tasks. Boys respond to competition, and so do girls. Contrary to previous literature (e.g., Datta Gupta et al. 2005, Gneezy et al. 2003, Gneezy and Rustichini 2004a) we also find that the gender of the opponent affects neither boys nor girls in any of the three tasks. The three performance measures we use here differ due to the difference in nature of the three tasks. This makes direct comparisons across tasks somewhat difficult, and we do find that there is actually an average decrease in performance when the children compete in dancing compared to the

[20] When we control for age in a tobit regression (upper limit 5 and lower limit 1), there is a gender difference in rating only for running, and age does not have a significant effect. It should be noted that the variation in age is very small. When controlling for age, boys and girls do not have different opinions concerning the rating of competition. It should be noted that the sample size is rather small.

individual performance, unlike in running and skipping rope. However, in each of the three tasks we find no gender difference in performance change, and this is our main result.

One possible explanation to the difference between our running result and that of Gneezy and Rustichini (2004a) is culture. It has previously been shown that culture affects important economic decisions such as labor market participation and fertility (e.g., Fernández and Fogli 2006), and the institutional setting has been found to influence competitive behavior (e.g., Balafoutas and Sutter 2010, Gneezy et al. 2009, Niederle and Yestrumskas 2008, Cotton et al. 2009, Niederle et al. 2009, Wozniak et al. 2010). For example, the gender gap in self-selection has been shown to disappear with performance feedback (Wozniak et al. 2010) and the difference in performance change vanishes with repetition of the competition (Cotton et al. 2009). Women have also been found to compete more than men in a matrilineal society whereas men compete more than women in a patriarchal society (Gneezy et al. 2009). Even though our study only includes children in Sweden, we can compare our running results to those of Gneezy and Rustichini (2004a).[21] Where we find no gender gap, Gneezy and Rustichini (2004a) instead find that among Israeli children only boys respond to competition in a running task. The specific mechanisms behind the different results in Sweden and Israel are unclear. It is possible that the more gender neutral culture in Sweden decreases the difference in competitive behavior between boys and girls in general, but also that it diminishes the degree to which tasks are gendered. If this is the case, this could explain why boys and girls compete equally in all tasks in our study.[22]

The results of two recent studies complicates this reasoning somewhat. Since we performed the study presented in this paper, there have been two other relevant studies. Sutter and Rützler (2010) look at willingness to compete among children aged 3 to 18 years old. Younger children are given the choice whether to compete or not in running 30 meters, whereas older children get the same choice for a math task. The authors find that boys are more competitive than girls in all age groups. When it comes to performance change in

[21] Even though the two studies differ somewhat in their design. In our experiment, we look at three different tasks, not only running. Moreover, the children compete in running 60 meters, which differs from the 40 meters used in Gneezy and Rustichini (2004a). However, the setups are similar in many aspects: both setups explore competitiveness as the performance change when running against someone versus running alone, the children were not aware of participating in an experiment, the teachers administered the running task, the matching procedure of the competing pairs was the same, and there was only intrinsic motivation for winning. We also included a control group, as in Gneezy and Rustichini (2004a).

[22] Children in Sweden do not receive grades until year 8 (age 14) thus a higher motivation for both boys and girls to perform well due to grade concerns is not a plausible explanation to why boys and girls compete equally.

running and math, there is however no gender gap. In an even more recent study, Cárdenas et al. (2010) explore the gender gap in competitiveness and risk taking among 9-12 year olds in Colombia and Sweden. Boys and girls are equally competitive in all tasks and all measures in Colombia (including running), whereas the results in Sweden are mixed, with some indication of girls being more competitive than boys in skipping rope and math in terms of performance change, whereas boys are more likely to choose to compete in general.[23] Cárdenas et al. (2010) also find that boys in both countries are more risk taking than girls, with a smaller gender gap in Sweden.

The absence of a gender gap in performance change in running in Austria and Colombia is surprising given the results in Israel. Both of these countries typically score as Israel on gender equality indices.[24] However, there are differences between the setup in Gneezy and Rustichini (2004a) and those in Cárdenas et al. (2010) and Sutter and Rützler (2010). An interesting avenue for future research would be to identify the specific components in explaining differences in the gender gap in competitiveness across a large number of countries using the exact same measures.

Making inferences about adult behavior from findings on children is not straightforward. Even though we do not find a gender gap among children in Sweden, it may be that male and female behavior change differently over time. Observing gender diversity in behavior among adults does not tell us the underlying reasons for these gender differences. For example, if a gender gap in behavior occurs during the teenage years, this could be caused by socialization or by the hormone surge that puberty brings along. More cross-cultural research and work on biological variables should also be of great interest. Thus far, studies looking at the importance of sex hormones to explain individual differences in competitiveness get mixed and inconclusive results. A study looking at competitiveness among men finds no relationship between self-selection into a tournament and current testosterone levels (Apicella et al. 2010). Buser (2009) finds that women are less likely to self-select into a tournament when progesterone and estrogen levels are high whereas Wozniak et al. (2010) find the opposite with women in the low-hormone phase being less competitive.[25] Meanwhile,

[23] The difference between our results and those of Cárdenas et al. (2010) in skipping rope is perhaps due to the larger sample size in the latter study (520 children).

[24] The Global Gender Gap Report 2009 ranks Austria as 42nd and Colombia as 56th out of 133 countries on gender equality.

[25] Apicella et al. (2010) find that neither circulating testosterone, facial masculinity (considered a proxy of hormone exposure during puberty), nor digit ratios (considered a proxy of prenatal hormone exposure) correlate

Zethraeus et al. (2009) find that exogenously providing estrogen or testosterone to women does not affect their economic preferences, though the authors do not look at competitiveness specifically.[26] More work is thus needed to disentangle the importance of sex hormones in explaining gender differences in competitiveness and other economic preferences.

Our findings open up interesting directions for further research. If competitive behavior among boys and girls is cultural and/or task dependent, competitive behavior should be studied in a variety of tasks and cultural settings. Since we find no gender differences among children in Sweden, it would also be of great interest to see if there is a gender gap in competitiveness among Swedish adults, and if so at what age this first occurs. It would also be interesting to in future studies collect information about the cultural background of the participants in experiments, to explore cultural variation in that sense too. Moreover, we do not use any extrinsic incentives in this study. An interesting extension would be to test the robustness of our results to extrinsic rewards such as money or e.g. pens.[27] Once we have answers to these questions it will be possible to make more general claims about gender and competitiveness, and possibly how and if this relates to labor market outcomes.

Acknowledgements

We are grateful for comments from Johan Almenberg, Anne Boschini, Hannah Riley Bowles, Tore Ellingsen, Armin Falk, Uri Gneezy, Moshe Hoffman, Magnus Johannesson, Astri Muren, Paul Nystedt, Mats Persson, David G. Rand and seminar participants at the Gender in Negotiation and Decision Making Seminar at Harvard Kennedy School, Harvard Business School, Harvard University, Ratio Institute, Santa Fe Institute, Stockholm School of Economics, Stockholm University, the 3rd Nordic Conference on Behavioral and Experimental Economics, University of Bonn and University of Zürich. Financial support from the Jan Wallander and Tom Hedelius Foundation (A.D., E.R.) and the Carl Silfvén

significantly with competitiveness in a sample of 98 young men. Moreover, Buser (2009) finds no effect of the cycle on competitiveness as measured by reaction to competition or risk preferences. This latter result contradict two studies that in turn also get opposing results when looking at competitive bidding/risk preferences. The first study finds that men and women who are menstruating (thus have low estrogen levels) act similarly (Chen et al. 2005), whereas a follow-up study finds that women menstruating or in the premenstrual part of the cycle act significantly different from men (Pearson and Schipper 2009).

[26] However, it could be the case that it is the long-term organizational effects of hormonal exposure that matter and not the effects from short-term exposure.

[27] Sutter and Rützler (2010) reward the children extrinsically when competing in running in Austria. Since they find no gender difference in performance change in running with this type of reward, this suggests that the lack of extrinsic reward is not necessarily what drives our results.

Foundation (E.R) is gratefully acknowledged, as well as financial support from the Swedish Council for Working Life and Social Research (FAS).

References

Apicella, C.L., Dreber, A., Gray, P.B., Hoffman, M., Little, A.C., Campbell, B.C. (2010). Androgens and competitiveness in men. Forthcoming in Journal of Neuroscience, Psychology and Economics.

Balafoutas, L., Sutter, M. (2010). Gender, competition and the efficiency of policy intervention. Working Papers in Economics 450, Göteborg University, Department of Economics.

Booth, A., Nolen P. (2009). Choosing to compete: The role of single-sex education. Mimeo.

Buser, T. (2009). The impact of female sex hormones on competitiveness. Mimeo.

Cárdenas, J.C., Dreber, A., von Essen, E., Ranehill, E. (2010). Gender differences in competitiveness and risk taking: comparing children in Colombia and Sweden. Forthcoming in Journal of Economic Behavior and Organization.

Ceci, S., Williams, W. (2006). Why aren't more women in science. American Psychological Association.

Chen, Y., Katuscak, P., Ozdenoren, E. (2005). Why can't a woman bid more like a man? Mimeo.

Cotton, C., McIntyre, F., Price, J. (2009). Gender differences disappear with exposure to competition. Working Paper 2010-11, University of Miami, Department of Economics.

Datta Gupta, N., Poulsen, A., Villeval, M.-C. (2005). Male and female competitive behavior – experimental evidence. IZA Discussion Paper, No. 1833.

Fernández, R., Fogli, A. (2006). Fertility: The role of culture and family experience. Journal of the European Economic Association, 4, 552-561.

Gneezy, U., Rustichini, A. (2004a). Gender and competition at young age. American Economic Review Papers and Proceedings, 118, 1049-1074.

Gneezy, U., Rustichini, A. (2004b). Executives versus teachers. Mimeo.

Gneezy, U., Leonard, K.L., List, J.A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. Econometrica, 77, 909-931.

Gneezy, U., Niederle, M., Rustichini, A. (2003). Performance in competitive environments: Gender differences. Quarterly Journal of Economics, 118, 1049 – 1074.

Günther, C., Ekinici, N.A., Schwieren, C., Strobel, M. (2009). Women can't jump? – An experiment on competitive attitudes and stereotype threat. Working Paper, University of Graz

Henschel-Pellet, H.A. (2001). Physical activity gender role stereotyping: An analysis of Children's Literature. Research Quarterly, 72.

Inzlicht, M., Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. Journal of Educational Psychology, 95, 796-805.

Niederle, M., Vesterlund L. (2007). Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics, 122, 1067-1101.

Niederle, M., Segal, C., Vesterlund, L. (2009). How costly is diversity? Affirmative action in light of gender differences in competitiveness. Mimeo.

Niederle, M., Yestrumskas, A.H. (2008). Gender differences in seeking challenges: The role of institutions. Mimeo.

O'Brien, L.T., Crandall, C.S. (2003). Stereotype threat and arousal: Effects on women's math performance. Personality and Social Psychology Bulletin, 29, 782-789.

Pearson, M., Schipper, B.C. (2009). Menstrual cycle and competitive bidding, MPRA Paper.

Price, J. (2008). Gender difference in the response to competition. Industrial and Labor Relations Review, 61, 320-333.

Shih, M., Pittinsky, T.L., Ambady, N. (1999). Stereotype susceptibility: Shifts in quantitative performance from sociocultural identification. Psychological Science, 10, 81-84.

Steele, C.M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. American Psychologist, 52, 613–629.

Sutter, M., Rützler, D. (2010). Gender differences in competition emerge early in life." IZA Discussion Paper 5015.

Wozniak, D., Harbaugh, W., Mayr, U. (2010). Choices about competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback. MPRA Paper.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Lindén Hirschberg, A., Johannesson, M. (2009). A randomized trial of the effect of estrogen and testosterone on economic behavior. Proceedings of the National Acadamy of Sciences of the United States of America, 106, 6535-6538.