



HAL
open science

Women can't jump?-An experiment on competitive attitudes and stereotype threat

Christina Günther, Neslihan Arslan Ekinici, Christiane Schwierén, Martin Strobel

► **To cite this version:**

Christina Günther, Neslihan Arslan Ekinici, Christiane Schwierén, Martin Strobel. Women can't jump?-An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior and Organization*, 2010, 75 (3), pp.395. 10.1016/j.jebo.2010.05.003 . hal-00849415

HAL Id: hal-00849415

<https://hal.science/hal-00849415>

Submitted on 31 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Title: Women can't jump?—An experiment on competitive attitudes and stereotype threat

Authors: Christina Günther, Neslihan Arslan Ekinci, Christiane Schwierien, Martin Strobel



PII: S0167-2681(10)00085-5
DOI: doi:10.1016/j.jebo.2010.05.003
Reference: JEBO 2538

To appear in: *Journal of Economic Behavior & Organization*

Received date: 25-1-2008
Revised date: 5-5-2010
Accepted date: 10-5-2010

Please cite this article as: Günther, C., Ekinci, N.A., Schwierien, C., Strobel, M., Women can't jump?—An experiment on competitive attitudes and stereotype threat, *Journal of Economic Behavior and Organization* (2008), doi:10.1016/j.jebo.2010.05.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Women can't jump? –

An experiment on competitive attitudes and stereotype threat

Christina Günther^a, Neslihan Arslan Ekinçi^b, Christiane Schwieren^{c,*}, Martin Strobel^d

^a*MPI of Economics, Evolutionary Economics Group, Kahlaische Straße 10, 07745 Jena, Germany*

^b*(r)evolution GmbH, Tulpenfeld 2, 53113 Bonn, Germany*

^c*University of Heidelberg, Department of Economics AWI, Campus Bergheim, Bergheimer Str. 58, 69115*

Heidelberg, Germany

^d*Universiteit Maastricht, Department of Economics, Faculty of Economics and Business Administration, P.O.*

Box 616, 6200 MD Maastricht, The Netherlands.

Abstract:

Gneezy et al. (2003) offer a partial explanation for the wage gap between men and women. In an experiment they found that women react less to competitive incentives. The task they used in their experiment can however be considered a male task. We replicate the experiment and extend it by treatments with a gender neutral task and a female task. For the male task we replicate their results, but for the neutral task women react as strongly to incentives than men and for the female task women react stronger than men. Our findings suggest a stereotype threat explanation. Women tend not to compete with men in areas where they (rightly or wrongly) think that they will lose anyway – and the same holds for men, although to a lower extent.

* Corresponding author. *Phone: +49 (0)6221 542953, Fax: +49 (0)6221 543630. e-mail: christiane.schwieren@awi.uni-heidelberg.de*

Keywords: Gender differences, competition, stereotype threat

JEL-Codes: J16, C91

Accepted Manuscript

Women can't jump? –

An experiment on competitive attitudes and stereotype threat

Abstract:

Gneezy et al. (2003) offer a partial explanation for the wage gap between men and women. In an experiment, they found that women react less to competitive incentives. The task they used in their experiment can however be considered a male task. We replicate the experiment and extend it by treatments with a gender neutral task and a female task. For the male task we replicate their results, but for the neutral task women react as strongly to incentives as men and for the female task women react slightly stronger than men. Our findings suggest a stereotype threat explanation. Women tend not to compete with men in areas where they (rightly or wrongly) think that they will lose anyway – and the same holds for men, although to a lower extent.

Keywords: Gender differences, competition, stereotype threat

JEL-Codes: J16, C91

1. Introduction

The gender wage gap and the limited success of women in reaching top positions in management cause continuous and emotional discussions in academia as well as in every-day life. Various explanations ranging from discrimination (e.g. Blau & Kahn 1994) to women's negotiation behaviour (e.g. Babcock & Laschever 2003) have been offered. Recently a new strand of literature (Gneezy et al. 2003, Gneezy & Rustichini 2004, Niederle & Vesterlund 2007) has added another possibility: Women compete less than men and even less when competing against men. Moreover women, when given the choice, prefer less competitive environments.

An initial benchmark was set by Gneezy et al. (2003). Students at the Technion in Haifa had to solve maze games in groups of six. Each student had to solve as many mazes as possible within a given time span. The experiment contained several treatments which varied in the sex composition of the group and the payoff schemes for the participants. In the Piece Rate treatment (baseline treatment) each group was composed of three men and three women. Subjects were given a fixed amount of money per solved maze. The same composition of groups holds for the Mixed Tournament and the Random Pay treatment. There however, only one participant received a compensation of six times the piece rate. In the Mixed Tournament treatment the amount was given to the subject with the highest performance. In the Random Pay treatment the amount was given to a randomly drawn participant.

The results show that men perform slightly better than women in the non-competitive treatments (Piece Rate and Random Pay). However, more importantly the performance of men increases significantly between the non-competitive treatments and the competitive treatment (Mixed Tournament) while women's performance does not¹. The results look different when looking at a fourth treatment with the same payoff scheme as in the Mixed

¹ Gender differences in risk taking are unlikely to explain the result, because the comparison between Random Pay and the Mixed Tournament treatment controls for risk preferences as far as this is possible.

Tournament treatment, but with the group composed of a single sex only (Single Sex Tournament). Still, women perform worse than men, but in the Single Sex Tournament treatment women increase their performance just like men do.

In another study, Gneezy & Rustichini (2004) look at competitive attitudes of children at young age when performing a running task. While the performance of boys and girls was roughly equal when time on a 40m sprint was measured individually, when they were competing with each other by running in pairs, it increased significantly with the boys but not with the girls. Girls did not increase their performance even when competing in single sex pairs.

The results were further qualified by Niederle & Vesterlund (2007) who focussed their investigation on the question of self-selection in tournaments. In a clever design they measured performance in a summation task and participants' willingness to submit this performance to a tournament payment scheme or not. Their main result is that men and woman perform equally well in the piece rate payment scheme and also equally increase their performance in a tournament payment scheme. However, if subjects are allowed to select the payment scheme, significantly more men than woman select the tournament scheme.

Our paper contributes to this literature by looking at an important factor, which has so far been neglected: the task in which subjects compete. We assume that women compete less in tasks that are perceived as typically male and therefore, women are stereotypically expected to perform worse than men, while they do not compete less than men in gender-neutral tasks or female tasks.

We base our hypothesis on the psychological literature on stereotype threat (e.g. Steele 1997, Steele & Aronson 1995, Ryan & Ryan 2005, Stangor et al. 1998, Keller 2007, Inzlicht & Ben-Zeev 2000). The concept of stereotype threat basically means that the activation of a specific stereotype may negatively impact task performance of the negatively stereotyped group. Most

of the research in this field exemplary uses the performance of females and black Americans in math tests as well as other tests of academic performance. The literature mentions many different potential mechanisms contributing to stereotype threat: arousal, anxiety, over-motivation, distracting thoughts, diminished self-efficacy or reduced performance expectations. Important mediating factors are task difficulty and identification with the task or with achievement goals. The relative importance and the interaction of these factors are still open to debate.

However, some predictions clearly arise from this research: Firstly, very subtle manipulations suffice to evoke salient stereotype threats. For example, just having to write down one's gender or race on the exam on the first page, or being clearly in the minority of one's sex within a group provokes saliency of the stereotype that can lead to stereotype threat. Secondly, stereotype threat impairs performance on tasks that are stereotypically related with lower performance of one group, especially for subjects in this group who generally perform well or above average in these tasks. Finally, stereotype threat seems to impair feelings of self-efficacy, and, following Stangor et al. (1998), performance expectations².

The contribution of our research therefore is to qualify the result that women compete less with respect to characteristics of the situation and the task. To test our assumptions, we first replicated the experiment by Gneezy et al. (2003). We then used the same paradigm, but with a sex-neutral task and a female task. Results confirm our hypotheses: In the male task, we replicate Gneezy et al.'s findings that women react less strongly than men to competitive incentives. For the neutral and the female tasks however, this does not hold. To the contrary: in the female task women improve their performance significantly while men do not.

In the remainder of the paper Sections 2 and 3 describe our experimental design and the procedure. Section 4 reports the results and Section 5 concludes with a discussion.

² For a contrary opinion see Ryan & Ryan (2005).

2. Experimental design

To test our hypotheses, we used a 3x2-design (three tasks and two incentive schemes). The three tasks we call the male task, the neutral task and the female task.

In the male task, participants were asked to solve maze games. The games were solved with paper and pencil, tracking a pencil through mazes presented on individual sheets. The participants had 15 minutes to solve a maximum of 20 mazes. The order of the mazes was fixed and the same for all participants. They received one point for each correctly solved maze. Points were transformed into payments depending on the incentive condition described below.

In the neutral task, the experimenters showed a slide with a letter on it. Within a limited amount of time participants had to generate and write down as many words as possible that started with this letter. Each word could be used only once, i.e., it was not possible to use the adjective to a verb or use different forms of the same verb. Furthermore, names were not allowed. Participants received one point for each correct word, and these points were transformed into money depending on the incentive condition.

The female task consisted of two subtasks, one a pattern matching (task A) and the other one a memory performance (task B). Both were taken from two different versions of the “WILDE” intelligence test (Jäger & Althoff 1983; Kersting et. al. in press). In the beginning, subjects had to read short descriptions of a workshop-setting (e.g., hotel name, presentation tools available). Later they had to answer question to this setting out of their memory. Between reading and answering questions subjects had to do a distraction task. They saw rows of three sketched faces and had to identify the face which was different from the two others in each row. Both tasks were done under time pressure, each correct answer yielded a point and finally the points from the two subtasks were added.

Why do we call these tasks male, neutral and female? Research on gender differences could establish only a few tasks were real differences in the performance of men and women could be observed (see e.g. Kimura 1999, 2004). But, stereotypes about what women and men are good at, do exist. Women are in general slightly less good and thought to be significantly less good at tasks that require spatial resolution, like rotating objects or navigating their way through a route. An example of such a task is solving mazes, our male task.

Even though women are stereotypically thought to be better than men in all tasks requiring verbal abilities, this could not be proven. Kimura has shown that women outperform men only in some categories of verbal abilities, namely in verbal memory and verbal association capabilities, but not in verbal fluency, which is required for our neutral task. For word generation women do not generally outperform men, but the task definitely lacks the negative performance stereotype against women.

In standard intelligence tests, women do not only perform better in verbal memory tasks, but even more so in tasks that require fast pattern recognition. Stereotypes with respect to these two kinds of tasks seem not particularly strong, but informal surveys among colleagues showed that both men and women indeed believe women to be better in both of these tasks.

The two incentive conditions of our experiment were either competitive or random pay.

In the competitive condition, participants were organized in groups of six. They were instructed about the specific task they had to do. Further, the instructions explained that the person who performed best on the task in the whole group would be the winner of this group. The winner would get an amount of € 0.50 per solved maze, € 0.20 per generated word, and € 0.20 per scored point respectively,³ in addition to the show up fee of € 3.-. The rest of the group would receive the show up fee of € 3.- only.

³ The incentives were designed in such a way that the ex-ante expected payoffs were similar across tasks.

In the random pay condition, participants were also organized into groups of six, but who won depended on pure luck: this was determined by throw of dice. Participants were shown the dice beforehand and it was insured by the experimenters that participants knew that performance in the task had no influence on their chances to be the winner. Payment was as in the competitive treatment – the person declared winner received performance based payment plus the show-up fee, while all others received the show-up fee of € 3,- only. By using a random pay scheme instead of a piece rate payment scheme we reduced the influence of risk attitudes as far as possible (see Gneezy et al. 2003). Still, subjects believing in heterogeneous abilities face different expected payoffs across different payment schemes. For example, a subject considering herself as a bad performer will consider her expected payoff lower in the competitive than in the random treatment. This is however inherent to the problem statement and can not be fixed unless we have a precise account of subjects' beliefs.⁴

In the following we will name our treatments: male competitive (MC), male random (MR), neutral competitive (NC), neutral random (NR), female competitive (FC) and female random (FR).

3. Experimental Procedure

The experiment was conducted at the Universitat Autònoma and the Universitat Pompeu Fabra in Barcelona, Spain. Subjects were recruited in sessions with 12 participants (6 male and 6 female) without specific focus on study background. In total 234 subjects participated (48 for each treatment MC⁵, C, and NR; 42 for treatment MR;⁶ 24 for each treatment FC and FR). The sessions were run as pen and paper experiments in large classrooms. Participants were seated separately from each other. After they had taken their seats, they received written

⁴ As it will turn out later we find effects for women and men, which cannot solely be reconciled with the simple explanation that either gender is more risk averse than the other.

⁵ One subject of treatment MC had to be excluded from the analysis because of drug consumption, which was recognized only after the experiment during the payment procedure. This decision was made before the analysis.

⁶ Due to some non-shows we had to run one session with 6 subjects only.

instructions which were also read aloud by the experimenter. Questions could be asked privately before the experiment started.

4. Experimental Results

In order to allow a comparison of the different tasks we report relative values of performance rather than absolute values. Therefore we divided the performance of each subject by the average performance of all subjects who did the same task, regardless of whether they played the competitive or the random payment scheme. A performance of $x\%$ indicates that the subject scored $x\%$ of the average number of points in a task. Throughout the section we use Mann-Whitney-U tests to investigate differences in distributions. Results do not qualitatively change if we investigate differences in means with t-tests instead; significance and non-significance of all results persist.

For the male task (maze games) we replicate the results of Gneezy et al. (2003). In MR women performed only at a level of 89.6% as compared to the average, but not significantly less than men who performed at a level of 101.7% (MWU exact, 1-tailed, $p = .188$). In MC women performed at a level of 90.5%, while men performed at 120.1%, significantly better than women (MWU exact, 1-tailed, $p = .018$).

Looking at women only across treatments MR and MC yield that women improve performance under competition but not significantly (MWU exact, 1-tailed, $p = .448$). Men improve across treatments MR and MC although the result is only marginally significant (MWU exact, 1-tailed, $p = .077$). We conclude:

Result 1: *In the male task, men do increase their performance in reaction to competitive pressure, but women do not.*

The picture changes when we look at the neutral task. In NR women outperform men with 91.0% in contrast to 82.4%. This is, however, not significant (MWU exact, 2-tailed, $p = .387$).

In NC the performance of men and women is nearly equal; it is 113.2% for women and 113.4% for men. The difference is not significant (MWU exact, 2-tailed, $p = .740$).

When looking at the increase of performance in the neutral task across treatments, we find that women increase their performance significantly (MWU exact, 1-tailed, $p = .018$) as well as men do (MWU exact, 1-tailed, $p = .009$). Hence, our second finding is:

Result 2: *In the neutral task, men as well as women increase their performance in reaction to competitive pressure.*

When looking at the female task, results are not that clear. There seems to be a tendency towards the opposite effect than in the male task, but due to the small sample size this cannot be finally established. In FR men outperform women slightly, with 98.3% compared to 89.8%. This difference is not significant (MWU exact, 1-tailed, $p = .123$). In contrast women outperform men with 108.9% compared to 103.0% in treatment FC. This is also not significant (MWU, exact, 1-tailed, $p = .299$). Interestingly, we find that men do not significantly increase their performance across treatments FR and FC while women do (MWU exact, 1-tailed, $p = .155$ for men, $p = .010$ for women). This could be a first indication that, when the negative stereotype threat for women is not only missing but even reversed, women engage into competition while men refrain from doing so.

A closer look at the data reveals that the result is driven by the distraction task. In treatment FR-distraction, men outperform women with 96.0% compared to 83.2%. This is marginally significant (MWU exact, 1-tailed, $p = .096$). In contrast, in treatment FC-distraction, women outperform men with 115.1% compared to 105.7% (not significant, MWU exact, 1-tailed, $p =$

.299). Like in the overall measure men did not significantly improve their performance while women did (MWU exact, 1-tailed, $p = .177$ for men, $p = .001$ for women).

Performances in the memory task are nearly identical across gender and payment scheme. While in treatment FR-memory men with 101.8% are marginally better than women with 100.0% (MWU exact, 1-tailed, $p = .507$), it is vice versa in treatment FC-memory. Women with 99.4% are marginally better than men with 98.8% (MWU exact, 1-tailed, $p = .392$). Both, men and women slightly decrease their performance across treatments FR-memory and FC-memory but not significantly (MWU exact, 1-tailed, $p = .383$ for men, $p = .318$ for women). Results of the female task can be summarized as:

Result 3: *In the female task, women but not men increase their performance in reaction to competitive pressure. The increase is driven by the distraction part (the simple recognition of sketched faces).*

Result 3 has to be taken with a grain of salt. While it comes out of the data, two issues have to be taken into account. Firstly, the female task was chosen because it was supposed to give women a better stand. However they perform worse than man in treatment FR (although insignificantly) and they only perform insignificantly better than men in FC. This makes alternative explanations for Result 3 plausible. If for example our FR sample of women was strongly biased towards bad performers, the significant improvement of women's performance would be natural.

Secondly, the effect is limited to the distraction task. This might be due to a ceiling effect for the memory task. The test-norms indicate an average performance in the norming sample of 13.3 remembered items on average for women and 11.4 for men, respectively. This coincides with the average absolute performance in our sample: 13.8. An alternative but related explanation is based on the dependence of the two subtasks. A high performance in the

distraction task might well lead to exhaustion⁷ and a bad performance for the memory task. This would also explain why the memory task is the only one in which subjects (men as well as women) show a lower performance under the competitive payment scheme compared to the random payment scheme.

Table 1 summarizes the performance results of all treatments as well as the results of the statistical tests.

<i>Task (av. absolute performance)</i>	<i>payment scheme</i>						<i>increase from random to competitive</i>	
	<i>random</i>			<i>competitive</i>			<i>women</i>	<i>men</i>
	<i>women</i>		<i>men</i>	<i>women</i>		<i>men</i>		
<i>male (10.7)</i>	89.6	$\leq^{1}_{.188}$	101.7	90.5	$\ll^{1}_{.018}$	120.1	$\leq^{1}_{.448}$	$<^{1}_{.077}$
<i>neutral (44.4)</i>	91.0	$\cong^{2}_{.387}$	82.4	113.2	$\leq^{2}_{.740}$	113.4	$\ll^{1}_{.018}$	$\ll^{1}_{.009}$
<i>female (35.2)</i>	89.8	$\leq^{1}_{.123}$	98.3	108.9	$\cong^{1}_{.299}$	103.0	$\ll^{1}_{.010}$	$\leq^{1}_{.155}$
<i>- distraction (21.4)</i>	83.2	$<^{1}_{.096}$	96.0	115.1	$\cong^{1}_{.299}$	105.7	$\ll^{1}_{.001}$	$\leq^{1}_{.177}$
<i>- memory (13.8)</i>	100.0	$\leq^{1}_{.507}$	101.8	99.4	$\cong^{1}_{.392}$	98.8	$\cong^{1}_{.318}$	$\cong^{1}_{.383}$

Table 1: Summary of results. The values are percentages with respect to the average performance for the corresponding task. Values in parentheses correspond to the average absolute score for the corresponding task. The symbol \leq means smaller than but not significant; $<$ means smaller than but only weakly significant; \ll means significantly smaller. Subscript indicates the significance level, while superscript indicates whether the test is one-tailed or two-tailed.

5. Discussion

In general a stereotype threat is supposed to lead to a decline in the performance compared to the usual performance of a subject. The introduction of a competitive payment scheme however, also results in an additional opposing effect, the general increase of performance (except for the memory task). A superposition of these two effects might therefore result in the absence of a performance decrease or even in an insignificant performance increase.

⁷ Even though one could think that pattern recognition is not effort-dependent, success in this task depends heavily on concentration and speed of performance, which in its turn is indeed effort-dependent. The memory task might actually be less subject to control by increased effort, which might be a third reason for not finding the effect in this task.

Our results confirm our assumption that stereotype threat plays a major role for the finding that women tend to avoid competition against men and improve performance to a lesser degree than men in situations of mixed gender competition. Even though most results on the topic reported so far are consistent with a stereotype threat explanation, to our knowledge none of the studies has analyzed this potential explanation systematically by varying the task and thus the related performance stereotype.

By doing exactly this we can show that it is not sex-composition per se that drives women's (or men's) reaction to competition, but that it plays an important role whether the task evokes a negative performance stereotype for women. We show that women do not improve performance in competition against men in a stereotypically male task. If however the task gives them the expectation of having an equal chance, they do improve performance in a competitive setting just as or more than men do.

There remains, however, one puzzle to be solved: Why do men in our neutral condition not react with reduced performance, even though there exists a negative performance stereotype for men with respect to verbal tasks? One potential reason could be the importance of the task for their self-esteem. Aronson et al. (1999) have shown that stereotype threat does impair performance especially for people who have a high interest in the task or who see the task as highly important for their self-esteem or esteem from others. In our student sample, performance in a verbal task is probably rather low in importance for male participants. In an academic setting, typically male skills (being analytic, good knowledge of math and technical issues) are often more important for esteem than the type of verbal skills needed in our experiment. If interest in and importance of the task is low, stereotype threat usually has no negative effect on performance.

This explanation is consistent with our results of our female tasks, where men do not react with a significant performance increase to the competition with women⁸. Pattern recognition could be considered an important skill for technical (and therefore male) professions and performance here might thus be considered more important for men than in a verbal fluency task.

The results of most other studies on gender differences in reaction to competition of which we are aware can be explained by stereotype threat. Most studies that find women to be less positively affected by competition than men use stereotypically male tasks (Datta Gupta et al. 2005, Gneezy et al. 2003, Niederle & Vesterlund 2007, Vandegrift et al. 2004⁹). For some of the other studies the link with stereotype threat is less clear, as more factors play a role in these studies. Stereotype threat might play a subordinate role among other factors in the studies by Price (2006), Gneezy & Rustichini (2004) and Antonovics et al. (2003).

What alternative explanations could explain our results? The explanation most of the other studies favour is a lower general tendency of women to compete. We do, however, not replicate this finding for a task that includes no negative stereotype for female performance and even find a tendency towards the contrary for a task where women are considered to perform better than men in general. Stereotype threat therefore seems to be the most applicable explanation for the findings reported here. In fact, a kind of double-stereotype threat is possible for competitive situations in stereotypically male tasks. Being competitive in itself is in all western cultures regarded as stereotypically rather male¹⁰, and in addition, being competitive in male settings for women still includes a negative stigma of being bitchy. The pure fact of being in a competitive situation against men might lead to stereotype threat if the

⁸ Women, in contrast, react with a significant increase in performance, but we do not want to over-interpret this for reasons discussed already in the previous section (after Result 3).

⁹ Vandegrift et al. (2004) even find that it is not sex alone that drives the avoidance of competitive situations, but skill in the task they use is very important at least in one of their competitive treatments.

¹⁰ See Gneezy et al. (2009) for an experiment in a society where this does not hold. As expected, the authors find that women compete more than men in a matrilinear society, where arguably competition is nothing stereotypically male.

performance stereotype of the task at hand is not explicitly favourable for women. A favourable performance stereotype has been shown to lead to stereotype boost in some situations (Shih et al. 2002) – which might explain the large positive effect of competition on men when competing in a stereotypically male task against women that has been reported in the literature.

With the focus on direct competition between the sexes, our study also contributes to the literature on stereotype threat. The psychological literature on stereotype threat usually uses specific ability tests, such as math-tests, which are stereotypically associated with good or bad performance of a specific group. The studies then compare absolute performance of women compared to men in these tests, which are performed individually and not as a direct competition between members of the two sexes. Saliency of the stereotype, perceived diagnosticity of the task for a specific ability and perceived susceptibility of the task to gender bias have been shown to affect whether stereotype threat affects performance or not.

Direct competition between the sexes might have a double effect. Apart from the above-mentioned direct effect through the stereotype of lower female ability to compete against men (without being perceived as bitchy), there may be an additional indirect effect. Having to compete against members of the opposite sex should make the sex-specific performance stereotype for the task at hand especially salient. This in turn should lead to a stronger effect of the stereotype threat on performance and might lead women who are conscious of this effect to refrain from competing against men (which is also in line with the results concerning choice of competitive situations by women, as reported in Niederle & Vesterlund 2007).

Our findings have practical implications for dealing with the under-representation of women in high managerial positions. As most of the contexts where women are under-represented are stereotypically associated with superior male achievement, they are classical situations for

stereotype threat. In addition, having to compete as such, which is stereotypically considered male behaviour, might be a situation of stereotype threat for women. Research on stereotype threat has shown that especially potential high achievers who strongly identify with the task at hand are susceptible to the negative effects of stereotype threat on performance (Aronson et al. 1999). Therefore, especially potentially high performing women are expected to perform worse than equally or less qualified men, or avoid competing against men altogether.

Our research suggests that one way to reduce the gender gap in wages and positions could be to teach women strategies to cope with the effects of stereotype threat or to design mechanisms that help to avoid stereotype threat from arising. This approach is already practiced in some symphony orchestras where job applicants play behind a curtain to avoid that judges know the gender of the player. Goldin & Rouse (2000) have shown that this practice indeed enhances hiring of female musicians. Given this result, comparable techniques could potentially not only reduce discrimination by employers, but also reduce the effect of stereotype threat on women, leading to better performance of potentially high achieving women and more possibilities for them to reach top positions.

Acknowledgements

We want to thank Jordi Brandts for help with the data collection, Lise Vesterlund for a helpful discussion and an anonymous referee for very good comments that helped to improve the paper.

References

- Antonovics K., Arcidiacono P. & Walsh R., 2003. Competing Against the Opposite Sex, Working Paper.
- Aronson J., Lustina M.J., Good C., Keough K., Steele C.M. & Brown J., 1999. When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat, *Journal of Experimental Social Psychology*, 35, 29-46.
- Babcock L. & Laschever S., 2003. *Women Don't Ask: Negotiation and the Gender Divide*, Princeton University Press, Princeton.
- Blau F.D. & Kahn L.M., 1994. Rising Wage Inequality and the U.S. Gender Gap, *American Economic Review*, 84, 23–28.
- Datta Gupta N., Poulsen A. & Villeval M.-C., 2005. Male and Female Competitive Behavior: Experimental Evidence, IZA Working Paper 1833.
- Gearry D.C., 1998. Male, Female - The Evolution of Human Sex Differences, American Psychological Association.
- Gneezy, U., Leonard, K.L. & List, J. (2009). Gender differences in competition. Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77, 1637-1664.
- Gneezy U., Niederle M. & Rustichini A., 2003. Performance in Competitive Environments: Gender Differences, *Quarterly Journal of Economics*, 118, 1049-1074.
- Gneezy U. & Rustichini A., 2004. Gender and competition at a young age, *American Economic Review Papers and Proceedings*, 114, 377-381.
- Goldin C. & Rouse C., 2000. Orchestrating Impartiality: The Impact of Blind Auditions on Female Musicians, *American Economic Review*, 90, 715-741.
- Inzlicht M. & Ben-Zeev T., 2000. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males, *Psychological Science*, 11, 365–371.
- Jäger, A.O., & Althoff K., 1983. *Wilde Intelligenztest (WIT)*, Göttingen, Germany: Hogrefe.

- Kersting, M., Althoff, K., & Jäger, A. O., 2008. WIT-2 Wilde Intelligenztest 2 [Wilde Intelligence Test: Completely revised version]. Göttingen, Germany: Hogrefe
- Keller J., 2007. Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance, *British Journal of Educational Psychology*, 77, 323-338.
- Kimura D., 1999. *Sex and Cognition*, MIT Press.
- Kimura D., 2004. Human sex differences in cognition; fact, not predicament, *Sexualities, Evolution & Gender*, 6, 45-53.
- Niederle M. & Vesterlund L., 2007. Do Women Shy Away from Competition? Do Men Compete Too Much?, *Quarterly Journal of Economics*, 122, 1067-1101.
- Price J., 2006. *Gender Differences in the Response to Competition*, Working Paper.
- Shih M., Ambady N., Richeson J.A., Fujita K. & Gray H.M., 2002. Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation, *Journal of Personality and Social Psychology*, 83, 638-647.
- Stangor Ch., Carr Ch. & Klang L., 1998. Activating Stereotypes Undermines Task Performance Expectations, *Journal of Personality and Social Psychology*, 75, 1191-1197.
- Steele C.M., 1997. A Threat in the Air. How Stereotypes Shape Intellectual Identity and Performance, *American Psychologist*, 52, 613-629.
- Steele C.M. & Aronson J., 1995. Stereotype threat and the intellectual test performance of African Americans, *Journal of Personality and Social Psychology*, 89, 797-811.
- Ryan K.E. & Ryan A.M., 2005. Psychological Processes Underlying Stereotype Threat and Standardized Math Test Performance, *Educational Psychologist*, 40, 53-63.
- Vandegrift D., Yavas A. & Brown P.M., 2004. Men, Women and Competition: An Experimental Test of Labor Market Behavior, Working Paper.