Gender differences in high profile jobs are a major factor in the gender gap in earnings.

**Several Possible Explanations:**

- Men and Women may differ in their abilities / preferences, which leads to occupational self selection.

- Discrimination, which leads to differential treatment of men and women with equal preferences and abilities.

  - There are field studies that support this view, e.g. Goldin and Rouse AER 2000, Wenneras and Wold Nature, 1997 <blind review… >.

Gneezy et al. propose and test an alternative (or supplementary) explanation:

**Women and men differ in their ability or propensity to perform in environments in which they have to compete against one another.**

Since incentive schemes commonly used in job evaluation and performance are often highly competitive, they may elicit different performance from men and women.

Little attention has been given to gender differences in competitive behavior in either the economics or the psychology literature.

**Goal:** Explore possible gender differences in behavior in competitive environments.

**Controlled Experiments:**

- Precisely measure performance.
- Exclude discrimination (and any expectation of discrimination.)
- No issue of selection of participants into different environments.
More general goals…

**Economic Relevance:**
- Understanding the reasons for the gender wage gap, the glass ceiling effect.
- Designing optimal incentive schemes.
- Deciding in unbiased ways, which are the most able persons.

**Policy Applications:**
- Affirmative action.
- Single-sex versus mixed schooling.

General experimental design:

- Real effort task (solving mazes)
- Different incentive schemes
- Different gender composition of groups
Experimental subjects:
Technion undergraduate students. (Degree in Engineering)
Each treatment: 10 groups of 6 participants each. In each treatment 30 women and 30 men.
384 participants in 64 experimental sessions.
Always different participants in different treatment: (Between subject design.)

Payment: Participants receive 20 NIS show up fee. (4NIS=1$).

The Task:
Solving Mazes.
(http://games.yahoo.com/games/maze.html)

After all participants solved one maze of level 2, the final part of the instructions were distributed.
Performance of participants when their reward is independent of the performance of others. (Benchmark)

**Treatment 1: Piece Rate Payment**
3 women and 3 men in the lab.
15 minutes to solve mazes (of difficulty level 2).
Participant receives 2 shekels for each solved maze.
Participants do not know how much the others earned.

Male average: 11.23
Female Average: 9.73.
Comparing the two distributions: The p-value of the Wilcoxon Mann-Whitney test is 0.2023, the difference is not significant.
Impact of Competition on the performance of men and women.

Treatment 2: Competitive Pay / Tournament
3 Women and 3 Men solve mazes for 15 minutes.
The person that solves the most mazes receives 12 shekels for each maze solved.
Others receive nothing.
As in the other experiments: Participants do not know how much others earned, i.e. also not who won.

Results:

Significant Increase in Performance
Tournament average: 12.95
Piece Rate average: 10.48
p-values of WMW-test: 0.007 significant differences.
**Significant Gender Difference**

Mean for men: 15, Women: 10.8. p-value of WMW test is 0.0004.

**Men strongly react to Tournament incentives:**

Tournament: 15

Piece Rate: 11.23: significant difference (p-value 0.001).

**Women do not react to Tournament incentives:**

Tournament: 10.8

Piece Rate: 9.73: no significant difference (p-value 0.6226).
Results so far:

- No significant gender gap in mean performance in the piece rate treatment: 1.5.
- Large and significant gender gap in mean performance in tournaments 4.2.

Furthermore, tournaments significantly increase the gender gap in mean performance compared to a noncompetitive piece rate: 4.2 >> 1.5.

Reasons for this gender gap: Tournament incentives
- do not increase mean performance of women.
- significantly increase mean performance of men.

What are some hypotheses that could account for this?
Gender Differences in Risk-Aversion?

Differences between tournament and piece rate:
- Payment depends on the performance of the other participants.
- Payment is uncertain.

Is the gender gap in mean performance in tournaments driven by the uncertainty only, through gender differences in risk-aversion?


To discern effect of risk aversion, need to consider incentives where the payment is uncertain, though independent of the performance of others.

**Treatment 3: Random Pay**
Group: 3 Men and 3 Women: Solve mazes for 15 minutes. At the end: One person is chosen randomly and receives 12 shekels for each maze she or he has solved. Other participants receive no payment additional to show-up fee.
Results Random Pay:

Mean for males: 11.83, for females: 10.33.
WMW p-value: 0.165. Difference is not significant.

**Random Pay versus Piece Rate**
Differences is not significant for men (0.6449) and women (0.6130).

**Random Pay versus Tournaments**
Difference is significant for men (0.0065) but not for women (0.6226).

So performance differences in mixed tournaments are not driven by the uncertainty, through gender differences in risk aversion.
Why do tournaments result in a significant increase in the gender gap in mean performance as compared to the piece rate?

Why do women not increase their performance in mixed tournaments while men do?

Some hypotheses:

1. Women do not compete against men:
   
   **Rational Explanation:**
   Participants have some ability to solve mazes. Output is jointly determined by effort, ability and noise. If it is common knowledge that women are (slightly) worse at solving mazes than men, their optimal effort level may be lower than that of men.

   **Additional Explanation:**
   Women may think that they are worse than men in solving mazes in competitive environments, beyond possible actual gender differences.

   …

   These explanations hinge on the identity of the competitors. Women might still be effective in competitive environments and be motivated by competitive incentives.
2. **Women do not compete at all:**

*Women cannot solve more mazes.*

*Women can not / do not want to compete per se.*

These are reasons that determine the behavior of women in tournaments *per se.*

(i) Women are not sensitive to incentive schemes at all.

(ii) Women do not like to compete.

* Different socialization
* Not increasing performance (effort) in the tournament as opposed to random pay: Contributing to a public good.

…

3. **The performance of men is too high (not that of women too low):**

Maybe men provide too much effort, and hence output, in tournaments, not women too little.

To investigate these issues: Single-sex tournaments: 6 Women only and 6 Men only.

Reason to have include Men single-sex tournament.

- Maybe men only compete a lot when women are around: Evolutionary Argument.

- If men only perceive other men as “real” competitors, men in mixed tournaments only compete against 2 other subjects. However, optimal effort in tournaments depends on the number of competitors. Hence we do not want to compare the performance of women in single-sex tournaments to the one of men in mixed tournaments.
Treatment 4: Single-sex Tournaments
Exactly like the other tournament, only each group of subjects now consists either of 6 women or 6 men. 5 sessions of each gender, i.e. 30 women and 30 men.

Results:
Men in single-sex tournaments look like men in mixed tournaments:
Men single sex: 14.3  Men mixed: 15 mazes
p-value of WMW: 0.5630: no significant difference.

Hence men are not strongly affected by the fact that they do not compete against women.

Are Women Competitive?
Women in Single-sex Tournaments Versus women in non competitive environments

Single-sex tournaments: 12.6
Random pay: 10.33  (p-value two-sided: 0.0469)
Piece Rate: 9.73: (p-value two-sided: 0.0148)

Women react strongly to tournament incentives in single-sex groups.
Women’s performance in mixed versus single sex tournaments:

Women in Single-sex Versus Mixed Tournaments:

Single-sex tournaments 12.6
Mixed Tournaments 10.8

p-value of WMW (two sided) is 0.1025.
Women versus Men in single-sex tournaments

Men: 14.3
Women: 12.6
WMW test: p-value 0.1346, the difference is not significant.
Gender gap in mean performance:

Mixed Tournament: 4.2
Single sex tournament: 1.7
Piece Rate: 1.5
Random pay: 1.5

Moving from mixed to single sex tournaments significantly reduces the gender gap in mean performance:

Repeat our bootstrap procedure, and find p-value of 0.082, hence the reduction in the gender gap is significant.

However no significant difference in the gender difference in performance when moving from single-sex tournaments to piece rate (0.459) and random pay (0.535).
The conclusions might be summarized as:

- Women do not compete against men.
- Women competing against women respond to competition as much as men do.
Apart from performance differences, 
Women may feel less confident in their ability and competence.

Do men and women feel equally competent and confident in their ability of solving mazes?

Do men and Women make different choices when they can chose the difficulty level in which they perform and will be evaluated?

**Treatment 6: Choice of Difficulty**
3 women and 3 men. 
Participants choose difficulty level from 1 = easy to 5 = hard and play in this level for 15 minutes. 
For mazes of level x receive x shekels for each maze solved. 
Participants do not know choices or performance of others.
The mean choice
Men: 3.4
Women: 2.6 (p-value 0.0065), the difference is significant.

Both arguments: That subjects take gender as a signal for ability and that women feel less confident in their ability than men are possible explanations.
Impact of Incentive Schemes

**Impact of Tournament incentives**

Mean performance in mixed tournaments: 12.95
Mean performance in single-sex tournaments: 13.47
Difference is not significant (p-value 0.62)

Overall performance is certainly not adversely affected by running single-sex as opposed to mixed tournaments.

Performance in tournaments: mixed and single sex: is significantly higher than under random pay and piece rate.

Tournaments are not only used to provide incentives, but also to determine very high performing participants.
Winner of the 10 tournaments in each treatment:

**Mixed Tournaments**: 7 men, 2 women (and 1 man and 1 woman tied in one group), average performance: 19.4
Average performance (1000 simulated tournaments): 19.26

**Single-sex tournaments**: By design: 5 women and 5 men. Average performance: 20.5: No significant difference (p-value: 0.52)
Average performance (1000 simulated tournaments):
Women: 19.59, Men: 21.04  Average: 20.32

Average performance of winners in mixed and single sex tournaments is not significantly different.

Running single-sex as opposed to mixed tournaments had no adverse effect on average performance and performance of the winner.
Running single-sex as opposed to mixed tournaments is a strong form of affirmative action: Quota setting. In this case: Came at no cost.
Experimental methods can thus tell us about group differences (majority/minority, men/women, also adults/children…)

**Economic Behavior of Children:**
http://harbaugh.uoregon.edu/index.htm


We study the development of bargaining behavior in children age 7 through 18, using ultimatum and dictator games. We find bargaining behavior changes substantially with age and that most of this change appears to be related to changes in preferences for fairness, rather than bargaining ability. Younger children make smaller dictator proposals than older children, and they also make and accept smaller ultimatum proposals. Even young children seem to be quite strategic in their behavior. Boys claim to be more aggressive bargainers than girls do, but they are not. We also find a relative height effect: within each experimental group, taller children make much smaller dictator offers. Since gender and height are correlated, height alone explains part of the gender effects. We argue that the existence of systematic differences in bargaining behavior across age supports the argument that culture is a determinant of economic behavior, and suggests that people acquire this culture during childhood. We argue that the height differences indicate that forces other than culture are also important.

In this paper we examine how risk attitudes change with age. We present participants from age 5 to 64 with choices between simple gambles and the expected value of the gambles. The gambles are over both gains and losses, and vary in the probability of the non-zero payoff. Surprisingly, we find that many participants are risk seeking when faced with high-probability prospects over gains and risk averse when faced with small-probability prospects. Over losses we find the exact opposite. Children’s choices are consistent with the underweighting of low-probability events and the overweighting of high-probability ones. This tendency diminishes with age, and on average adults appear to use the objective probability when evaluating risky prospects.

Economic experiments that you can perform at home on your children. Kate Krause and William T. Harbaugh. 1998 Working paper.

This paper describes some simple economic experiments that can be performed and analyzed by undergraduates, using children as subjects. Experiments on children are particularly well suited for undergraduate research projects. The analysis of experimental data is generally straightforward. Since little is currently known about the economic behavior of children,
even rather standard experiments protocols can produce new results. We argue that by conducting experiments on children economists can gain insight into the origins of preferences, the development of bargaining behavior and rationality, and into the origins of “irrational” behavior in adults. This paper give examples of successful research projects and discusses what they reveal about the economic behavior of children. We also include suggestions for new projects, and a discussion of practical aspects of experimenting on children, such as appropriate incentives and human subjects approval. Protocols are included that can be used to replicate our experiments.


In this chapter we examined trusting and trustworthy behavior in children. We asked whether our approach to these complex tasks develop with age. We conducted a set of experiments with students in third, sixth, ninth and twelfth grade to investigate this question. Like adults, the students in our experiment did not follow the game-theoretic prediction of passing and returning zero tokens, and we conclude that even the 8-year-old children show clear evidence of trusting and
trustworthy behavior. In contrast to common perception we do not find that the youngest participants are more trusting than older cohorts, nor do we find that they are much less trustworthy.


In this paper we examine the extent to which consumption choices by 7 and 11-year-old children and college undergraduates satisfy the axioms of revealed preference. We find that choices by even the 7-year-olds are considerably more likely to obey revealed preference axioms than would be true if they were choosing randomly. 11-year-olds do better still, while college students do no better than 11-year-old children. We also find that mathematical ability is not correlated with choosing rationally. We argue that this evidence suggests that the ability to choose rationally is not innate, but that it does develop quickly.


If preferences are defined over consumption levels then there should be no difference between the compensation demanded to give up a market good and the willingness to pay for the same good. However many experimental studies have found large disparities between these measures. While some argue that this should be seen as
strong evidence for reference-dependent preferences others argue that it is nothing more than a mistake. If it is a mistake, we would expect it to be smaller among more experienced traders. To obtain large differences in market experience to test this, we compare the endowment effect in five to ten-year-old children with that of undergraduates. We find that very large increases in age and experience do not reduce the apparent endowment effect. We take this as strong evidence in support of the hypothesis that the endowment effect reflects underlying preferences and is not a mistake.


In economic experiments on the voluntary provision of public goods, adults are more altruistic than the narrow definition of self interest would predict. In this paper we see if the same is true among six- through twelve-year-old children. We find that in aggregate the level of altruistic behavior of children is similar to that of adults, but that in repeated experiments the pattern over time is different. Younger children’s contributions tend to increase in later rounds of the experiments, while the contributions of older children, like those of adults, tend to decline. We also find that demographic and family variables do not explain differences in altruistic tastes across children, but a measure of group attachment does. Finally, contributions in a subsequent dictator experiment are correlated with first-round contributions.
in the public good experiment, but are not strongly correlated with last-round contributions.