Exp Econ (2016) 19:269–280
DOI 10.1007/s10683-015-9440-5

ORIGINAL PAPER

Incentives for creativity

Sanjiv Erat[1] · Uri Gneezy[1,2]

Received: 16 December 2013 / Revised: 23 March 2015 / Accepted: 21 April 2015 /
Published online: 3 May 2015

Abstract We investigate whether piece-rate and competitive incentives affect creativity, and if so, how the incentive effect depends on the form of the incentives. We find that while both piece-rate and competitive incentives lead to greater effort relative to a base-line with no incentives, neither type of incentives improve creativity relative to the base-line. More interestingly, we find that competitive incentives are in fact counter-productive in that they reduce creativity relative to base-line condition. In line with previous literature, we find that competitive conditions affect men and women differently: whereas women perform worse under competition than the base-line condition, men do not.

Keywords Creativity · Competition · Intrinsic motivation · Extrinsic motivation · Choking · Behavioral economics

Electronic supplementary material The online version of this article (doi:10.1007/s10683-015-9440-5) contains supplementary material, which is available to authorized users.

✉ Sanjiv Erat
serat@ucsd.edu

Uri Gneezy
ugneezy@ucsd.edu

1 Rady School of Management, University of California, San Diego, USA

2 CREED, University of Amsterdam, Amsterdam, Netherlands

Competition is for horses, not artists.
Bela Bartok

1 Introduction

Creativity is central to many activities, including entrepreneurship and research and development at large corporations. Can incentives be used to improve the creative process? If you are in charge of choosing a compensation scheme for a worker who needs to come up with a creative solution to a problem, should you compensate the worker based on her creative performance? And if so, what type of incentives work best? In this paper, we compare the effect of different incentives for being creative. We start with perhaps the simplest form of incentives, piece-rate incentives based on creativity ratings, and we test whether creativity improves when one introduces such incentives relative to simply asking people to be creative without offering them any extrinsic incentives.

Next, we consider competitive incentives, which offer a more convex link between performance and reward than piece-rate incentives, and consequently might have a stronger effect on effort levels (Lazear and Rosen 1981), and thus on (creative) performance. Moreover, because the evaluation of creativity might have a large subjective component, a risk-averse participant might prefer competitive incentives (and put in greater effort and show higher performance) because in such an incentive scheme it is the difference in participants' performances that determines the reward.

On the flip side, while high-powered competition incentives might encourage greater effort, the same incentives also have the potential to cause choking under pressure (Ariely et al. 2009), according to which people may have a harder time coming up with a creative idea when they know they are competing with another person.

Our key results are that incentives did not improve the creative output relative to the case in which participants are not offered any external monetary incentives for creativity. Moreover, the type of incentives matter, and competitive incentives reduce creativity relative to piece-rate incentives.

A possible explanation for these results is that differences in effort, as measured by time, could explain the differences in creativity between treatments. To test for this explanation we also manipulated how much time participants have to come up with a creative solution (either limited at 10 min or not limited). The effect of changing the time did not interact with the above mentioned results.

These results also relate to the discussion of the relation between incentives and creativity in psychology. In particular, the "crowding-out hypothesis" in creativity argues that although intrinsic motivation enhances creativity, extrinsic motivation may actively undermine a person's intrinsic motivation, and thus reduce overall creativity. Therefore, external incentives such as monetary rewards and competition may lower creativity. A large body of evidence supports the crowding-out phenomenon in multiple other domains (e.g., Deci and Rayen 1985; Gneezy and Rustichini 2000; Frey and Jegen 2001). However, relatively few studies have investigated the effect of monetary rewards and competition on creativity, and to our knowledge, none of these studies have compared different forms of extrinsic incentives to examine the relative benefits of using them to improve creativity.

Early studies in psychology have found some evidence to support the hypothesis that monetary incentives may crowd out motivation and thus reduce creativity (e.g., Glucksberg 1962; Amabile 1982; Amabile et al. 1986). However, more recent studies have challenged the finding that financial rewards are detrimental to creativity (e.g., Conti et al. 2001; Eisenberger et al. 1998; Eisenberger and Rhoades 2001; Eisenberger and Shanock 2003).

The studies in psychology and management that do find a negative relationship between monetary rewards and creativity have typically used fairly small monetary rewards (and often employed non-standard subjects). For instance, Amabile's (1982) study raffled off three prizes to the children who made the best collages at a party. As Gneezy and Rustichini (2000) noted, although extrinsic incentives may crowd out certain behaviors, increasing the size of the extrinsic incentives can more than offset this effect.

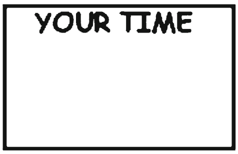
The rest of the article is organized as follows: Sect. 2 elaborates on the experimental design and the specific incentive treatments. Section 3 presents the results and explores alternative explanations. Finally, Sect. 4 closes with the key implications of our findings.

2 Experimental design

Our experimental design employed a rebus puzzle as the creative task. A rebus is a puzzle made with words and/or pictures with a hidden and non-obvious solution. Figure 1 below shows two examples of such puzzles with the solutions being "your time is up" and "2 (too) funny 4 (for) words." We use the task of designing a puzzle as our creative task.

Our between-participant design had six different treatments. In all treatments, participants were told they could draw as many puzzles as they wished on the sheets of paper handed to them, but they would have to mark one as the final submission when they handed it in.

We use a 3 (no incentives vs. piece rate vs. competition) × 2 design (10-min time limit vs. no time limit). In the two no-incentives treatments, BASE-L and BASE-NL, while we informed the subjects in these treatments that we valued creativity, we did not give them any additional incentives and only told them that a



Funny Funny
Words Words Words Words

Fig. 1 Example rebus puzzles

panel of judges would be evaluating their submissions on a scale from 1 to 10 for creativity. These two no-incentives treatments differ only in the amount of time the participants had to design their puzzles, with BASE-L participants having 10 min, and BASE-NL participants having no time constraint from the practical perspective (a maximum of 1 h).

In both the piece-rate treatments, PR-L and PR-NL, we told the participants that a panel of judges would be evaluating their submissions on a scale from 1 to 10 for creativity, and that the average score they obtained would determine their payment—receiving $2 per each point. The two treatments using piece rates differ in the amount of time the participants had to design their puzzles. In the PR-L, participants had 10 min to work, and in the PR-NL a maximum of 1 h.

In the competition treatments, we told the participants that judges would be evaluating their submissions on a scale from 1 to 10 for creativity (just as in the piece-rate treatment), and that we would then match them with another participant. Payment was based on the average score they received from the judges and how this average score compared to the other participant's average score. If they scored higher than the participant with whom they were matched, they received $4 per each point of creativity score; if the scores were the same, they received $2 per each point of creativity score; if their score was lower, they received no payment for this task. As in the piece-rate treatments, we had two competition treatments in which participants either had 10 min to work on the task (C-L) or had a 1 h time limit (C-NL).

Full instructions for the participants in the rebus design task are given in Appendix 1 in Supplementary Material.

We conducted the experiment in a laboratory setting. The 257 participants in the experiment received a participation fee of $5, in addition to any potential earnings from the experiment. Table 1 shows the number of participants in each of the conditions. When participants arrived at the lab, the experimenter randomly assigned them to one of the conditions. The experimenter also noted the participant's gender, as well as the time at which the participant received the instruction sheet and the time at which the participant turned in his/her entry. The time information allows us to determine how long the participant worked on the task.

Once all participants had turned in their entries, we employed eight independent raters; four saw these entries in a specific order, and the remaining four saw them in the reverse order. The instructions to the raters and the submissions the raters evaluated are given in Appendix 2 in Supplementary Material. These independent raters received a $20 flat fee for judging the creativity of the entries on a scale of 1–10.

Table 1 Number of participants

	10 min	Unlimited time
Base	45	46
Competition	41	43
Piece-rate	41	41

Because we had asked the judges to "rate the creativity of puzzles (i.e., give a rating to each puzzle based on whether it has original ideas, is innovative, and is clever)," these ratings are obviously subjective assessments. Given our interest in examining the effects of incentives on creativity, verifying that the creativity ratings that the judges assign reflect some (underlying) objective metric is important. We did this by looking at the correlation between the ratings the eight different raters awarded, and found an average correlation of 0.23.[1] As a more concrete single statistical test that also allows us to eliminate some of the subjectivity in creativity ratings, we looked at the correlation between the average score assigned by the four raters who saw the puzzles in one order and the average score assigned by the four raters who saw the puzzles in reverse order. This correlation turns out to be positive and significant ($r = 0.57$, $t = 11.2$, $df = 255$, $p < 0.001$).

Next, to obtain the average creativity score for each of the submitted rebus puzzles, we averaged the scores the eight raters assigned.[2]

3 Results

Figure 2 shows the average creativity scores and the standard errors of submitted entries in each of the treatments.

We first test the hypothesis that extrinsic incentives may crowd out creativity. Table 2 gives the regression results of creativity score against a indicator variables representing the specific incentive conditions, and another indicator variable representing unlimited time versus limited time.

Model 1 in Table 2 shows that relative to base-line case, while piece-rate does not significantly change ($\beta = -0.32$, $t = -1.51$, $p = 0.13$), the decrease in creativity in the competitive incentives treatment is significant ($\beta = -0.69$, $t = -3.31$, $p < 0.001$).

Result 1: Competition incentive reduces creativity relative to the no-incentives treatment.

Result 2: Piece-rate incentive does not change creativity relative to the no-incentives treatment.

The above results are important because they show that at best extrinsic monetary incentives are ineffective, and at worst, they might backfire and actually result in lowering creativity.[3]

[1] Out of the $8 \times 7/2 = 28$ between rater correlations, 23 were positive and significant at least within the 10 % confidence level, and the remaining 5 were positive but not significant.

[2] Alternate ways of aggregating the creativity scores, including normalizing each individual rater's scores, and then averaging these normalized ratings yields identical results.

[3] We thank a reviewer for pointing out that the participants do not get feedback in the no incentive conditions (compared to the incentive conditions where participants can use the payment to infer their own score), and that this could reduce motivation. Thus, it is interesting that despite this possible lower motivation, the creative output in the no incentive condition might be higher compared to the incentive conditions.

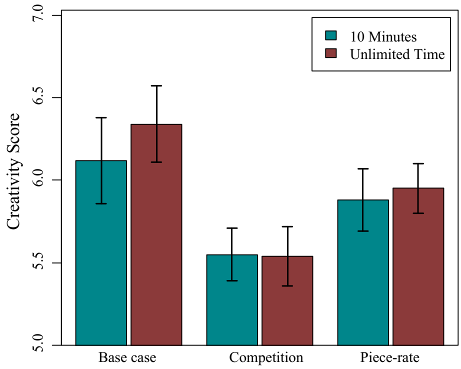


Fig. 2 Average creativity and standard errors for the 6 conditions

Table 2 Regression results comparing no-incentive and incentive conditions

Variable	Model 1 Coefficient (standard error)	Model 2 Coefficient (standard error)	Model 3 Coefficient (standard error)
Const	6.18*** (0.19)	5.86*** (0.15)	6.12*** (0.26)
1{Base}	–	0.32 (0.21)	–
1{Competition}	−0.69*** (0.21)	−0.37** (0.17)	−0.57* (0.30)
1{Piece Rate}	−0.32 (0.21)	–	−0.24 (0.32)
1{Unlimited Time}	0.10 (0.16)	0.10 (0.16)	0.23 (0.34)
1{Base}:1{Unlimited Time}	–	–	–
1{Competition}:1{Unlimited Time}	–	–	−0.23 (0.42)
1{Piece Rate}:1{Unlimited Time}	–	–	−0.15 (0.42)
R-squared	0.047	0.047	0.047
No. observations	257	257	257
Sigma	1.307, df = 253	1.307, df = 253	1.311, df = 251

*** $p < 0.001$; ** $p < 0.05$; * $p < 0.1$

The heterogeneity in how people react to competitive versus piece rate incentives prompted us to also conduct a second regression, summarized in Model 2 of Table 2, in which we found that competitive incentives marginally decrease creativity relative to piece rate incentives ($\beta = -0.37, t = -2.17, p = 0.03$). Thus, when the incentives were based on competition, our participants were less creative compared to those that relied on piece-rate incentives.[4]

[4] In the competition condition, only the winners learn their creativity score. We thank a reviewer for pointing out that this could lead to lower motivation in the competition treatment relative to the piece-rate one, and could be one explanation for the difference in creativity scores.

Result 3: Piece-rate incentives increase creativity relative to competitive incentives.

As argued earlier, because the incentives in the competition are steeper, one might expect the participants to invest more effort with competitive incentives relative to the piece-rate treatment. Although participants may work as hard as they can during the 10 min we allocate to them, giving them unlimited time to work on the task might override this difference in creativity in the two extrinsic incentive conditions. In particular, participants in the competitive treatment may choose to stay longer and put more effort into coming up with a creative solution. To test for this hypothesis, we conducted the regression (summarized in Model 3 of Table 2) with an interaction term for the time constraint and the incentive condition.

Our results indicate that, irrespective of the incentive condition, we fail to find any significant difference in the creativity scores when participants were given unlimited time. This suggests that the 10 min the participants had to draw the puzzle was not a major constraint on obtaining a highly creative outcome. Equally importantly, if Result 1 was caused by participants being differently constrained by the 10-min time limit in one treatment versus the other, we should expect the significant difference between the competitive-incentives and base-line to disappear when participants have unlimited time. However, this decrease in creativity scores persists irrespective of the time constraints ($\beta = -0.57$, $t = -1.86$, $p = 0.06$).

All our previous results show that explicit monetary incentives reduce creativity. As argued earlier, there might be two reasons why this occurs: crowding out of motivation (Amabile 1982; Amabile et al. 1986), and choking under pressure (Ariely et al. 2009).

To understand the relative importance of these two mechanisms, we had also collected data on how much time the subjects spend on the rebus task. Table 3 shows the mean and median time spent in each of the conditions.

Using a Mann–Whitney test, we find that people spent significantly more time in piece rate compared to base-line condition ($W = 993.5$, $p = 0.04$); and people spent significantly more time in competition compared to base-line condition ($W = 1139$, $p = 0.09$). Thus, effort levels (as measured by time) increases with extrinsic incentives. This finding, of course, runs counter to the crowding out (of motivation) theory which would have predicted that participants should be less motivated and consequently should spend less time they are incentivized compared to the no-incentive condition.

Result 4: Participants spend more time on the task when there are extrinsic incentives compared to the case where there are no incentives.

The above result along with our previous results that creativity decreases with extrinsic incentives is consistent with choking.

Table 3 Time spent designing the rebus puzzles

Base-line	Mean = 13.8 (median = 13 min)
Competition	Mean = 17.2 (median = 15 min)
Piece rate	Mean = 19.2 (median = 17 min)

Table 4 Relationship between time spent and creativity score

Variable	Coefficient (standard error)
Const	6.09*** (0.29)
1{Competition}	−0.96*** (0.27)
1{Piece-rate}	−0.39 (0.28)
Time·Spent	0.02 (0.01)
R-squared	0.11
No. observations	121
Sigma	1.22, df = 117

*** $p < 0.001$;** $p < 0.05$;* $p < 0.1$

Table 5 Variance of creativity scores

	Unlimited time	10 min
Base	2.93	2.38
Competition	1.14	1.22
Piece-rate	1.50	0.97

Since the incentives did increase effort levels, it is useful to test whether greater effort levels (as measured by time spent on the task) are related to greater creativity. Table 4 shows that the time spent on the task is not significantly related to creativity score ($\beta = 0.02$, $t = 1.6$, $p = 0.11$). Moreover, the difference in creativity between the incentive conditions still persists even after we control for the time that people spent on the task.

Result 5: Allowing participants to spend more time on the task does not affect the difference between treatments or improve their creativity.

Although we found that piece rate is ineffective in increasing the average creativity whereas competition incentives in fact decrease creativity relative to base line, an alternate reason for our finding could be that participants are conditioning their strategy on the payment method. In particular, participants may choose to take more risks and come up with "crazier" solutions when competing, because competitive incentives may promote risk-taking behavior (Hvide 2002). We test this possibility by using an F-test to look for significant differences in variance of creativity scores in the treatments.

Table 5 below shows the variance of creativity scores in each of the treatments. The only pair-wise comparisons of variances that are (marginally) significant are between Base-line and Competition, and between Base-line and Piece-rate.[5]

[5] For the case of unlimited time, in the test between base-line and piece rate, the ratio of variances is 2.45 (F = 2.44, num df = 45, denom df = 40, p < 0.01). For unlimited time, in the test between base-line and competition, the ratio of variances is 1.95 (F = 1.95, num df = 45, denom df = 42, p = 0.03). For the case of limited time, in the test between base-line and piece rate, the ratio of variances is 1.96 (F = 1.95, num df = 44, denom df = 40, p = 0.03). For the case of limited time, in the test between base-line and competition, the ratio of variances is 2.56 (F = 2.56, num df = 44, denom df = 40, p < 0.01).

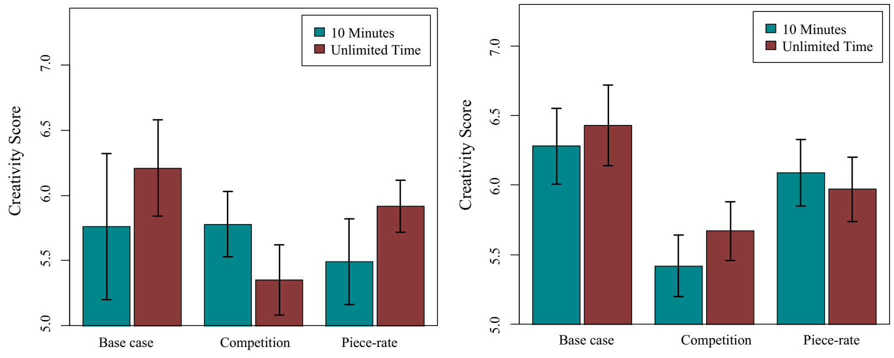


Fig. 3 Average creativity scores for men (*left frame*) and women (*right frame*)

Importantly, the variances of creativity scores in the incentive conditions are not significantly different.

The test of heterogeneity in effort levels combined with the test of heterogeneity in variances show that our results cannot be explained in terms of differences in effort levels or by participants actively pursuing a riskier strategy in the competitive conditions, and are consistent with the "choking hypothesis".

3.1 Gender differences

Past research found that men are more responsive to competitive incentives than women (Gneezy et al. 2003; see the survey in Croson and Gneezy 2009). This finding suggests the possibility of an interesting gender difference in the effect of competition that we found in our aggregate data. Figure 3 shows the average creativity scores for men and women for the treatments.

Our data indicates differences in how men and women react to these incentives. Table 6 reports a more detailed analysis, and shows the results of one regression comparing incentive condition against no-incentive (Model 1), and another regression comparing between the two incentive conditions (Model 2). Women's creativity is significantly lower in the competition condition relative to the base-line condition ($\beta = -0.8$, $t = -3.2$, $p = 0.001$), whereas this pattern is not statistically significant for men ($\beta = -0.8 + 0.34 = -0.46$, std err $= 0.33$, $t = -1.42$, $p = 0.16$). Also, women's creativity is significantly lower in the competition treatment relative to the piece-rate one ($\beta = -0.49$, $t = -2.2$, $p = 0.03$), whereas this pattern is not statistically significant for men ($\beta = -0.49 + 0.31 = -0.18$, std err $= 0.32$, $t = -0.55$, $p = 0.6$).

Result 6: Competition reduces creativity relative to piece-rate incentives and no-incentive condition for women, but not for men

This gender difference in the reaction to competition could have emerged from differences in effort. To address this possibility, we also checked for any possible difference in effort (time) between men and women.

Table 6 Regression results for gender differences in creativity scores

Variable	Model 1 Coefficient (standard error)	Model 2 Coefficient (standard error)
Const	6.35*** (0.20)	6.03*** (0.16)
1{Base}	–	0.31 (0.26)
1{Competition}	−0.80*** (0.25)	−0.49** (0.22)
1{Piece-rate}	−0.31 (0.26)	–
1{Male}	−0.34 (0.37)	−0.31 (0.25)
1{Base}:1{Male}	–	−0.03 (0.45)
1{Competition}:1{Male}	0.34 (0.45)	0.31 (0.35)
1{Piece-rate}:1{Male}	0.03 (0.45)	–
R-squared	0.05	0.05
No. observations	256	256
Sigma	1.31, df = 250	1.31, df = 250

*** $p < 0.001$; ** $p < 0.05$; * $p < 0.1$

Table 7 Regression results for gender differences in time spent

Variable	Coefficient (standard error)
Const	13.93*** (1.06)
1{Competition}	3.28 (2.09)
1{Piece-rate}	5.13 (3.32)
1{Male}	−0.32 (1.54)
1{Competition}:1{Male}	0.23 (3.51)
1{Piece-rate}:1{Male}	0.55 (3.51)
R-squared	0.06
No. observations	121
Sigma	8.92, df = 115

*** $p < 0.001$;** $p < 0.05$;* $p < 0.1$

Table 7 shows that, irrespective of the incentive condition, there is no significant difference between men and women in how much time they spent. Thus, if we use time as a proxy for effort, the gender differences we found in the effect of competition on creativity do not result from differences in effort levels.

The result regarding the interaction of gender and incentives with regard to creativity is important from a policy perspective. For example, Balafoutas and Sutter (2012) and Niederle et al. (2012) showed experimentally that affirmative action policies that target gender differences in competitiveness can increase the efficiency of the labor environment. Similar arguments can be made regarding policies that attempt to increase success on creative tasks: Incentives seem to affect men and women differently, and the relative success in performing the tasks depends on the specific type of incentives chosen.

4 Conclusion

Creativity is valuable in many endeavors. This study examines whether we can improve creativity through monetary incentives, and if so, what type of incentives work best. Past literature offers a mixed message on the effects of monetary incentives.

Our experimental design allows us to measure both the effort levels and the creativity, thus we can offer a more nuanced view of how extrinsic incentives affect creativity. Moreover, we compare competitive incentives with piece-rate incentives, and test how different types—and not merely the presence—of extrinsic incentives affect creativity.

Our main message is that while the incentives we used in the reported experiment increase the effort levels significantly, these same incentives do not improve creativity. In fact, competitive incentives reduced the creativity of our participants, a result mostly driven by the effect on women. The gender difference in the effect of competitive incentives is consistent with past literature showing that women react differently to competitive incentives than men.

The results in this paper provide an initial step in the study of different methods to incentivize creative performance. These results do not tell us directly when incentives might interact with the task type, and if these results might be equally applicable to non-creative and possibly routine tasks. Understanding this could be a fruitful line for future research. Moreover, within the domain of creative tasks, future research can expand the current study to other dimensions of creativity, such as R&D races, and other types of incentives, such as team compensation. Our message though is that the use of incentives in promoting creativity is not simple and predictable by standard assumptions.

References

Amabile, T. M. (1982). Children's artistic creativity: Detrimental effects of competition in a field setting. *Personality and Social Psychology Bulletin, 8*, 573–578.

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50*, 14–23.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies, 76*(2), 451–469.

Balafoutas, L., & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science, 335*, 6068.

Conti, R., Collins, M. A., & Picariello, M. L. (2001). The impact of competition on intrinsic motivation and creativity: Considering gender, gender segregation and gender role orientation. *Personality and Individual Differences, 30*, 1273–1289.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 1–27.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Eisenberger, R., Armeli, S., & Pretz, J. (1998). Can the promise of reward increase creativity? *Journal of Personality and Social Psychology, 74*, 704–714.

Eisenberger, R., & Rhoades, L. (2001). Incremental effects of reward on creativity. *Journal of Personality and Social Psychology, 81*, 728–741.

Eisenberger, R., & Shanock, L. (2003). Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation. *Creativity Research Journal, 15*, 121–130.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys, 15*(5), 589–611.

Glucksberg, S. (1962). The influence of strength of drive on functional fixedness and perceptual recognition. *Journal of Experimental Psychology, 63*, 36–41.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics, 118*(3), 1049–1074.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics, 115*(3), 791–810.

Hvide, H.K. (2002). Tournament rewards and risk taking. *Journal of Labor Economics*, *20*, 877–898.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy, 89*, 841–864.

Niederle, M., Segal, C., & Vesterlund, Lise. (2012). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science, 59*, 1–16.