Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Flipside of Comparative Payment Schemes

Thomas Buser, Anna Dreber

To cite this article:
Thomas Buser, Anna Dreber (2016) The Flipside of Comparative Payment Schemes. Management Science 62(9):2626-2638.
http://dx.doi.org/10.1287/mnsc.2015.2257

MANAGEMENT SCIENCE
Vol. 62, No. 9, September 2016, pp. 2626–2638
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2015.2257

The Flipside of Comparative Payment Schemes

Thomas Buser
School of Economics, University of Amsterdam, 1001 NJ Amsterdam, Netherlands; and Tinbergen Institute, 1082 MS Amsterdam, Netherlands, t.buser@uva.nl

Anna Dreber
Department of Economics, Stockholm School of Economics, SE-113 83 Stockholm, Sweden, anna.dreber@hhs.se

Comparative payment schemes and tournament-style promotion mechanisms are pervasive in the workplace. We test experimentally whether they have a negative impact on people's willingness to cooperate. Participants first perform in a simple task and then participate in a public goods game. The payment scheme for the task varies across treatment groups. Compared with a piece-rate scheme, individuals in a winner-takes-all competition are significantly less cooperative in the public goods game. A lottery treatment, where the winner is decided by luck, has the same effect. In a competition treatment with feedback, winners cooperate as little as participants in the other treatments, whereas losers cooperate even less. All three treatments lead to substantial losses in the realised social surplus from the public good while having no significant impact on performance. In a complementary experiment, we aim to shed light on the psychological mechanisms behind our results.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2015.2257.

Keywords: economics; behavior and behavioral decision making; game theory and bargaining theory; comparative pay; competition; cooperation; gender differences; incentive schemes

History: Received December 3, 2013; accepted May 17, 2015, by Teck-Hua Ho, behavioral economics. Published online in *Articles in Advance* November 23, 2015.

Introduction

Many companies use relative reward schemes whereby employees earn a bonus if they perform better than their colleagues. Moreover, hierarchical structures mean that in many organisations, employees find themselves in constant competition for promotions. This is meant to provide incentives for higher performance. Often, however, the success of a project depends not only on individual performance but also on the willingness of people within an organisation to cooperate with each other. Here, we ask whether comparative payment schemes have a detrimental psychological effect on people's willingness to cooperate. That is, we ask whether such schemes negatively affect an individual's willingness cooperate, even when cooperation does not affect the probability of earning the bonus.

To answer this question, we conduct an online experiment with a large subject pool in which participants are randomly allocated to perform a simple task under different incentive schemes. Following the task, we measure their willingness to cooperate in a public goods game with a randomly selected group of other participants whom they have not previously interacted with.

Participants in the control group are paid according to a piece rate whereas those in the "competition" treatment are paid according to a competitive winner-takes-all incentive scheme. In the "lottery" treatment, participants are paid according to a lottery scheme that reproduces the payoff variance, but not the performance dependence, of the competitive payment scheme: the winner is randomly chosen. Finally, in the "feedback" treatment, participants compete and are informed about the outcome of the competition before making their choice of how much to contribute to the public good.

The lottery treatment serves to establish whether the effect is the same when the outcome is due to luck rather than performance, which mimics comparative pay based on a (very) noisy performance measure. The feedback treatment eliminates the role of beliefs about the outcome of the competition. It also allows us to investigate whether feedback on the outcome, which is usually present in a workplace context, has an additional effect on cooperation.

An extensive literature finds that people are willing to cheat to artificially increase their own performance, lie about their performance, or sabotage the performance of others in order to win a competition (e.g., Lazear 1989, Konrad 2000, Chen 2003, Falk et al. 2008, Harbring and Irlenbusch 2011, Charness et al. 2013). Our focus is different. Namely, we are interested in whether there are behavioural spillovers from competition to other, unrelated settings, and decisions.

A couple of past studies have looked at spillovers from competitive situations to choices in the public goods game when both settings involve the same participants. Savikhin and Sheremeta (2013) find that simultaneous participation in a public goods game and a lottery contest decreases suboptimal overbidding in the contest, but public goods game contributions are not affected compared to when these games are played in isolation. Cason and Gangadharan (2013) instead find that simultaneously competing in a double-auction market lowers contributions in a threshold public goods game.

A number of further studies investigate spillovers of various forms of competition and rivalry to games and decisions other than contributions in the public goods game. Brandts et al. (2009) find that what they dub competitive rivalry—where player A chooses between players B and C to play a prisoner's dilemma—diminishes the rivals' altruistic disposition toward each other and player A, mainly because of negative emotions linked to loss of control and exclusion. Gill et al. (2013) find that participants in an experiment who are paid according to a lottery are slightly more likely to cheat vis-à-vis the experimenter compared with participants who are paid a fixed wage. Goette et al. (2012) conduct a lab-in-the-field experiment in the Swiss army using naturally occurring groups (army platoons). They find that between-group competition increases the incidence of antisocial punishment of out-group members. Balafoutas and Sutter (2012) find that affirmative action schemes favouring women in a competition do not affect efficiency in a subsequent coordination game relative to standard competition. Mollerstrom (2013), on the other hand, finds a negative effect of affirmative action, in general, on cooperation. Finally, in a field experiment, Bandiera et al. (2013) find that competitive incentives for fruit pickers change team composition (workers choose to work with others of a similar performance level rather than with friends) and lead to an overall productivity loss, probably because of a loss of social connections.

Closest to our design is Chen (2010), who also allocates participants to either a competitive or a piece-rate payment scheme and finds that competition increases nonutilitarian value choices in the moral trolley problem. After providing feedback, he also finds that competition losers but not winners are less likely to donate 10 cents to charity. An important difference between our design and Chen (2010) is that in our main treatment participants do not receive feedback on the competition outcome, allowing us to separate the effect of competitive incentives from the effect of feedback. In a subsequent paper, Ter Meer (2014) finds that tournament incentives increase deception in a sender-receiver game played with the same opponent, whereas there is no effect of cooperative incentives relative to piece rate.

Also related are a number of recent papers on the effects of market interactions on morality and social preferences. Falk and Szech (2013) let participants in a lab experiment decide between receiving money or saving the life of a "surplus" lab mouse. Participants are more likely to save the mouse when the decision is individual than when they make the decision through a multilateral market mechanism. Herz and Taubinsky (2013) show that exposure to market pressure serves as an anchor and lowers the minimum amount participants are willing to accept in an ultimatum game but not the amount they offer to others. Al-Ubaydli et al. (2013) find that priming participants using phrases related to markets and trade increases trust in a trust game through a positive effect on the beliefs about others' trustworthiness compared with a control group. Finally, Cappelen et al. (2013) find no significant effect of a market prime on honesty.

Our study adds to these previous findings in several ways by asking whether organisations face a trade-off between incentivising their workers and fostering a cooperative environment. In our design, the cooperative game involves a different set of participants from the task stage. This allows us to exclude reciprocity and punishment motivations as mechanisms. Also, the effects we measure differ from the effects uncovered in past experimental studies on sabotage and cheating because in our design, cooperating does not affect the chance of winning. Rather, our design measures a psychological spillover effect, which exists over and above the potentially perverse incentive effects of comparative payment schemes.

We find that participants in the competition treatment contribute significantly less in the public goods game compared with the control group, suggesting a negative effect of competition on the willingness to cooperate with others. Participants in the lottery treatment cooperate as little as those in the competition treatment, suggesting that the negative effect of comparative payment on the willingness to cooperate is present when the outcome is due to pure luck, and it is therefore not caused by participants having to perform under competitive pressure. In the feedback treatment, we find that winners cooperate as little as participants in the competition and lottery treatments, demonstrating that pessimistic beliefs about one's own performance cannot explain the competition effect. The feedback treatment's losers contribute less than the winners, and overall, participants in the feedback treatment contribute significantly less than participants in the competition and lottery treatments, showing an effect of feedback over and above the effect of comparative pay. All three treatments lead to

Table 1 **Experimental Treatments**

	Control	Competition	Lottery	Feedback
Comparison	Individual	Three randomly selected performances	Random groups of four	Three randomly selected performances
Payment	Four cents per point	16 cents per point if beat other performances, otherwise zero	Random group member gets 16 cents per point, others get zero	16 cents per point if beat other performances, otherwise zero
Feedback	No	No	No	Yes

substantial losses in terms of the realised social surplus compared with the piece-rate scheme while having no significant impact on performance in the task.

We then explore possible psychological mechanisms behind these results—in particular, cognitive load and priming with competitive norms. We use our experimental data to show that cognitive load is an unlikely explanation for our results. Using a supplementary online experiment, we show that priming participants with competitive norms using the scrambled sentences paradigm (Srull and Wyer 1979) has the same detrimental effect on public goods contributions as the competitive payment schemes. This suggests a mechanism whereby working under comparative payment schemes triggers competitive norms that continue to influence decisions in other settings. We also discuss other possible interpretations and remaining uncertainties.

Following a large literature that finds strong gender differences in the reaction to competitive incentives (Gneezy et al. 2003), in the reaction to competition outcomes (Buser 2014), and in the taste for competition (Gneezy and Rustichini 2005, Niederle and Vesterlund 2007), we also analyse whether the effect of competition on cooperation is different for men and women. We find that effects are generally stronger and more significant for the male subsample.

Experimental Design and Data Collection

Participants were recruited on Amazon's Mechanical Turk (MTurk) online labour market. MTurk is a platform where employers can recruit workers for short tasks in exchange for small payments (most tasks take less than five minutes with participants earning less than $1). MTurk has become increasingly popular with researchers across the social sciences as a platform for running incentivised experiments. MTurk makes it feasible to recruit much larger samples at a substantially lower cost and replication studies have found that despite the small amounts that are typically paid, results are very similar to those obtained in the lab (Horton et al. 2011, Amir et al. 2012).

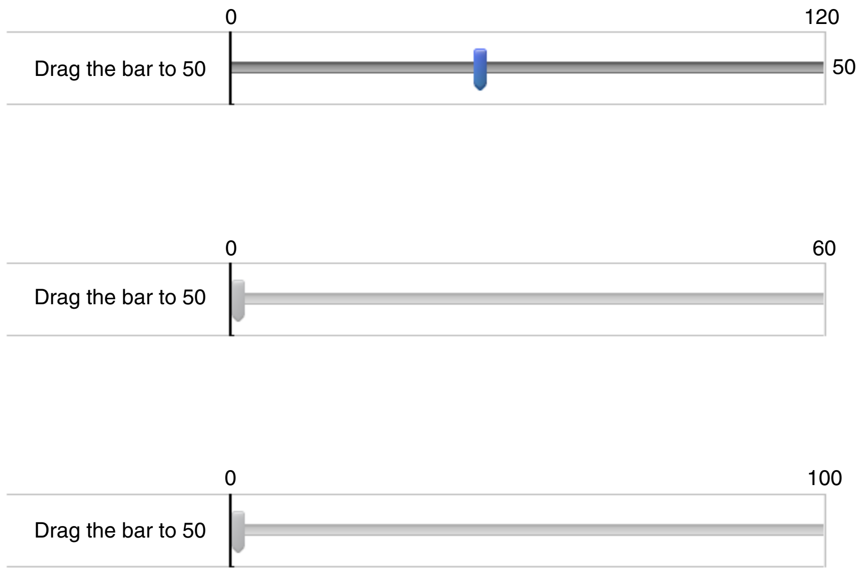
After accepting to participate in our study, participants on MTurk were referred to an external website (Qualtrics) and were randomised into one of four treatments, each consisting of two rounds. They were told that one of the two rounds would be randomly selected for payment. In the first round of all treatments, participants had to perform a simple task during two minutes. The treatments differ in the way participants were paid for their performance (see Table 1). Participants allocated to the control group were paid according to a piece rate of four cents per point. Those in the competition treatment had their performance compared with three other randomly selected performances and were paid 16 cents if they beat the others and zero otherwise, receiving no feedback on the outcome during the study. In the lottery treatment, one in four participants randomly received 16 cents per point, whereas the rest received nothing. Participants in the competition and lottery treatments did not receive feedback about their payment at any point during the experiment. Finally, the payment scheme of the feedback treatment is identical to the competition treatment, but participants received immediate feedback on whether they won or lost at the end of their performance. The choice of payment schemes is closely related to the design of Gneezy et al. (2003).

The task is an adapted version of the slider task of Gill and Prowse (2012). Participants faced a screen full of slider bars and had to position as many of them as possible on the 50 mark before time ran out (see Figure 1).[1] The main difference with Gill and Prowse (2012) is that the sliders are positioned in one neat row but that the end point varies from slider to slider. This task has the advantage that it is easy to implement in an online survey.

In the second round, all participants played a standard public goods game with three other randomly selected individuals (a different set of individuals from those they interacted with in the task round). These are randomly picked from the pool of other participants in the same treatment group. The participants in the three treatments were informed that these were not the same individuals they had interacted with in the first round. Participants received an allocation of 80 cents and had to decide how much to keep and how much to allocate to the group. Money allocated to the group was doubled and evenly divided

[1] In this task, it is obvious to the participants whether they have positioned the bars on the 50 mark or not.

Figure 1 (Color online) The Slider Task



among the four group members. The amount allocated to the group is our measure for the willingness to cooperate.

Finally, the participants answered a short questionnaire. We asked for gender and age, and we elicited risk attitudes through a simple, nonincentivized question: "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). Dohmen et al. (2011), using representative survey data from Germany, find that the question predicts both incentivised choices in a lottery task and risky behaviour across a number of contexts.

The average earnings in the sliders task across all treatments were \$1.34 and the average earnings for the public goods game were 1.25. Participants also received a fixed 50-cent fee for completing the experiment. The median amount of time it took participants to complete the experiment was seven minutes. Participation in the experiment was restricted to individuals living in the United States.

Data

Table 2 describes the data and sample. We have a total of 1,700 participants, 934 of whom are male and 766 of whom are female. They are on average 30 years old. The participants were randomly allocated to the four treatments with allocation to the feedback treatment being twice as likely to ensure a sufficient number of winners.

Our participants correctly positioned an average of 26.1 sliders with men performing significantly better than women ($p = 0.000$; rank-sum test). Contrary to the findings of some past studies on gender differences in cooperation in the public goods game, average allocations are exactly equal for men and women in our sample ($p = 0.899$; rank-sum test). But like many past studies, we find that men rate themselves significantly more risk seeking ($p = 0.000$; rank-sum test).[2]

Table 2 Data

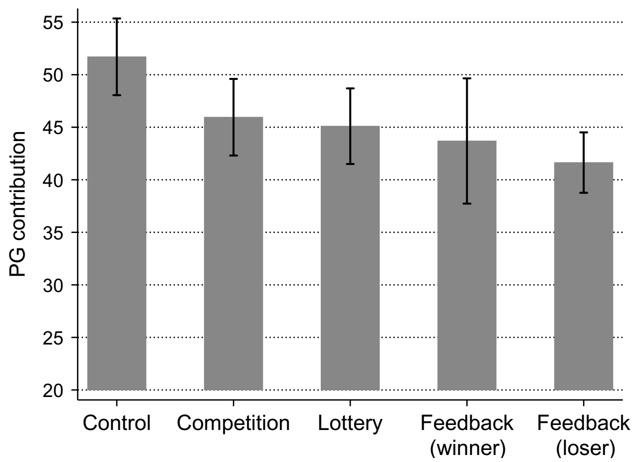
	Sample	Men	Women
Observations			
Control	334	173	161
Competition	351	184	167
Lottery	360	208	152
Feedback	655	369	286
Total	1,700	934	766
Variable means			
Performance	26.1	28.9	22.6
PG allocation	45.4	45.4	45.4
Age	30.1	28.3	32.3
Risk seeking	6.6	6.9	6.1

Note. PG, public goods.

A potential worry with online experiments is selective attrition. If many participants drop out after treatment is revealed and their decision to drop out is correlated with their preferences, this could lead to spurious differences between treatments. To avoid a large number of dropouts, we made it clear to participants that they would only receive their show-up fee if they completed the entire survey. Moreover, the average payment in our experiment was very

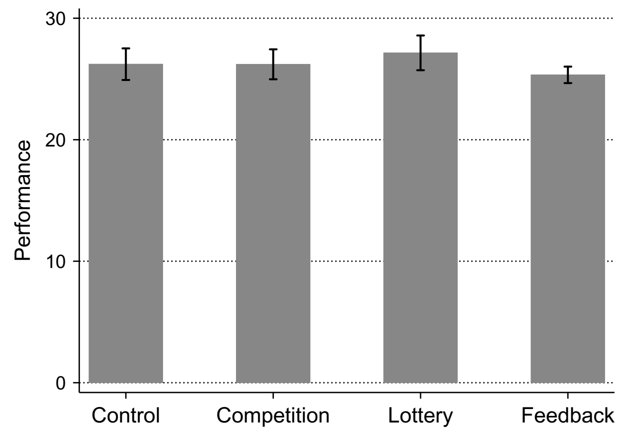
[2] See Croson and Gneezy (2009) for an overview of studies on gender differences in risk attitudes, competitiveness, and social preferences.

Figure 2 **Average Allocation by Treatment**



Note. The error bars show 95% confidence intervals. PG, public goods.

Figure 3 **Average Performance by Treatment**



Note. The error bars show 95% confidence intervals.

generous compared with the standards of MTurk, and incentives for completing the survey were correspondingly high.

Fifty-five participants dropped out of the survey, 25 of these after treatment was revealed (which is small compared to a sample of 1,700). These 25 observations are evenly distributed across the four treatments ($p = 0.258$; Fisher's exact test).[3] Selective attrition is therefore not an important concern for our study.

Results

Figure 2 shows average allocations to the group for each treatment. In each of the treatments, allocations are lower on average than in the control group. The nonparametric Kruskal–Wallis test returns $p < 0.001$, showing that allocations differ significantly across treatments.[4] Participants in the competition treatment, who compete against three other individuals, give 46 cents on average compared with 52 cents in the control group ($p = 0.025$; rank-sum test). Participants in the lottery treatment, who face random payoffs, give 45 cents, which is also significantly less than those in the control treatment ($p = 0.013$). Participants in the feedback treatment, who receive immediate feedback after competing, give 42 cents on average, which is less than both the control group ($p < 0.001$) and those in the competition treatment ($p = 0.084$). Focusing on the feedback treatment, we find that winners give 44 cents and losers give 42 cents. This difference between winners and losers is not significant ($p = 0.630$).

The lower allocations in the treatments lead to substantially lower social surplus compared with the control group (where we define social surplus as the difference between the average public goods game payoff and the guaranteed minimum average of 80 cents that participants would get if nobody contributed anything). The achieved surplus is 11% lower in the competition treatment and 13% lower in the lottery treatment compared with the control group. In the feedback treatment, the difference is 19%. Performance, on the other hand, hardly varies across treatments ($p = 0.707$; Kruskal–Wallis test). This is illustrated in Figure 3. This means that in our experimental setting, there is no positive impact of comparative incentives on performance, which could make up for the loss in social surplus in the public goods game.

In Table 3, we pursue these results more formally using ordinary least squares (OLS) regressions with controls for gender, performance, and age (column (2)) and risk attitudes (column (3)). We regress the public goods game contribution on treatment dummies with control participants being the omitted category. The results from the nonparametric tests are confirmed. Participants in all three treatments give significantly less than the control participants and the effects are robust to the inclusion of controls.[5]

Conditional on the full set of controls, participants in the competition and lottery treatments contribute approximately 12% less in the public goods game compared with participants in the control group. This indicates that comparative pay has a detrimental spillover effect on people's willingness to cooperate with others. This effect is present whether the outcome is decided by superior performance or by pure

[3] Of the 25 dropouts, 5 were allocated to the control treatment, 5 to the competition treatment, 9 to the lottery treatment, and 6 to the feedback treatment. Of the other 30 dropouts, none read the treatment-specific instructions.

[4] Note that all tests in the paper are two-sided.

[5] The results are qualitatively and quantitatively similar when we use Tobit regressions. We have also run ordered probit regressions and the conclusions remain the same.

Table 3 OLS Regressions (Dependent Variable: Public Goods Game Contribution)

	(1)	(2)	(3)
Competition	−5.748**	−5.725**	−5.156**
	(2.622)	(2.601)	(2.521)
Lottery	−6.602**	−6.274**	−5.239**
	(2.606)	(2.578)	(2.534)
Feedback (winner)	−8.010**	−6.264*	−7.059**
	(3.532)	(3.554)	(3.541)
Feedback (loser)	−10.064***	−11.295***	−10.941***
	(2.363)	(2.354)	(2.301)
Male		1.598	−0.841
		(1.718)	(1.695)
Performance		−0.372***	−0.307***
		(0.082)	(0.081)
Age		−0.107	−0.062
		(0.078)	(0.076)
Risk seeking			3.083***
			(0.323)
R^2	0.011	0.023	0.077
N	1,675	1,675	1,675

Note. Robust standard errors are in parentheses.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

Table 4 Differences in Public Goods Game Contributions Between Treatments (p-Values in Parentheses)

Competition − Feedback	3.904	4.548	4.998
	(0.086)	(0.047)	(0.023)
Lottery − Feedback	3.050	3.956	4.884
	(0.176)	(0.080)	(0.027)
Competition − Feedback (winner)	2.261	0.539	1.904
	(0.522)	(0.880)	(0.590)
Lottery − Feedback (winner)	1.407	−0.010	1.820
	(0.689)	(0.998)	(0.605)
Competition − Feedback (loser)	4.316	5.570	5.785
	(0.068)	(0.020)	(0.012)
Lottery − Feedback (loser)	3.462	5.021	5.702
	(0.140)	(0.034)	(0.014)
Feedback (winner) − Feedback (loser)	2.055	5.031	3.882
	(0.539)	(0.142)	(0.255)
Gender, performance, and age		√	√
Risk attitudes			√

Note. p-Values are from Wald tests of differences in public goods game contributions between treatments after OLS regressions corresponding to columns (1)–(3) in Table 3.

luck. This is interesting because in the workplace context, outcompeting others depends on performance and luck to varying degrees, depending on how well the organisation is able to monitor the true performance of its employees.

Strikingly, winners in the feedback treatment are no more generous, giving 15% less compared with the control group. This is interesting for several reasons. First, it shows that the effects of the competition and lottery treatments cannot be explained by participants trying to make up for an expected loss. Second, it shows that there is a negative effect of competitive environments even on the winners. Third, it shows that the negative effect of comparative pay persists after the payoff uncertainty has been resolved.

Losers in the feedback treatment are even less cooperative, contributing 19% less compared with control participants and 5% less compared with the winners. This is important because many competitions, such as, for example, the competition for promotions in a company, create only one winner but many losers. It also shows that when feedback is present, as it usually is in a workplace context, comparative pay has an even stronger negative effect on cooperation.

All the results reported above are robust to the inclusion of controls for gender, performance, age, and self-reported risk preferences. Gender and age are not significantly correlated with contributions to the public good. Performance in the slider task is significantly negatively correlated with contributions, which might be due to people who are more driven by financial rewards both putting in higher effort and being less willing to share. More risk-seeking individuals contribute significantly more, which might be due to higher willingness to face the psychological risk of contributing more than the other group members.

We explore the significance of the difference between the various treatments in Table 4, which reports p-values from Wald tests for the difference in public goods game contributions between the feedback treatment and the competition and lottery treatments. The tests confirm that, conditional on the full set of controls, participants in the feedback treatment give significantly less than participants in both the competition and the lottery treatments. The difference between the contribution of feedback losers and the competition and lottery treatments is even larger and always significant, whereas the contribution of feedback winners does not differ significantly from the competition and lottery treatments.

Table 4 also shows that the difference in contributions between winners and losers in the feedback treatment is not significant at conventional levels. However, success in a tournament is not randomly assigned but correlates positively with effort and talent. Winners and losers might therefore differ along various dimensions. Conditional on a participant's score, however, winning or losing only depends on the performance of the randomly assigned opponents and is therefore random (Buser 2014).

In Table 5, we explore the effect of feedback on public goods contributions when performance is properly controlled for. This is achieved by regressing public goods contributions on a loser dummy controlling for a third-order polynomial in the competition score using only participants in the feedback treatment. Losers give 6.5 cents less than winners. This

Table 5 Effect of Feedback (OLS Regressions; Dependent Variable: Public Goods Game Contribution)

	(1)	(2)	(3)
Feedback (loser)	−6.538*	−6.509*	−4.690
	(3.832)	(3.845)	(3.847)
Performance	−0.618	−0.502	−0.522
	(1.219)	(1.225)	(1.192)
Performance^2	0.006	0.005	0.014
	(0.040)	(0.040)	(0.039)
Performance^3	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)
Male		−0.259	−3.560
		(2.843)	(2.844)
Age		0.101	0.089
		(0.137)	(0.134)
Risk seeking			3.014***
			(0.508)
R^2	0.012	0.012	0.065
N	649	649	649

Notes. Robust standard errors are in parentheses. The sample consists of participants in the feedback treatment.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

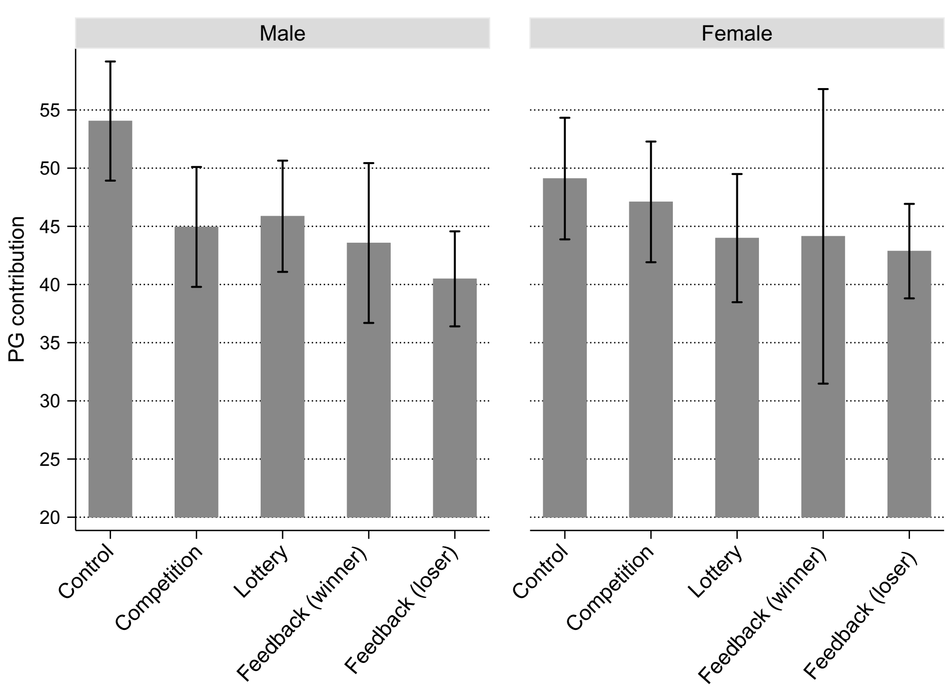
effect is significant at the 10% level. However, when a control for risk preferences is added, this shrinks to 4.7 cents and becomes statistically insignificant. These results suggest that negative feedback might have an additional causal effect on the willingness to cooperate with others. Negative emotions triggered by losing the competition are a plausible pathway for this effect. Past research demonstrates that emotions such as anger and sadness have a significant impact on choices in social preference games (Harlé and Sanfey 2007, Andrade and Ariely 2009).

Our results are largely in agreement with those of the studies that resemble our design most closely. Our finding of a negative impact of competitive incentives on cooperation is in accordance with Ter Meer (2014). Chen (2010) also finds that competition losers are less generous (albeit in a dictator game with a charity) but does not find a difference between competition winners and those who work under a piece rate.

Finally, we have a look at gender differences in the effect of comparative pay on the willingness to cooperate. Past research shows that the attitudes toward and reactions to competition vary between men and women, with men generally being more competitive. Figure 4 shows the differences in public goods contributions across treatments separately for men and for women. The differences are stronger for men. This is reflected by nonparametric tests: whereas contributions vary significantly across all treatments for men ($p = 0.001$; Kruskal–Wallis test), this is not the case for women ($p = 0.217$).

Table 6 shows regression results separately by gender. In the male subsample, participants in all treatments give less than the control group, whereas in the female subsample, only the difference between the feedback treatment losers and the control group is significant. However, the differences in the treatment effects for men and women are not generally significant at even the 10% level, with the exception

Figure 4 Average Allocation by Treatment (by Gender)



Note. The error bars show 95% confidence intervals. PG, public goods.

Table 6 OLS Regressions by Gender (Dependent Variable: Public Goods Game Contribution)

	(1) Male	(2) Female	p	(3) Male	(4) Female	p	(5) Male	(6) Female	p
Competition	−9.101**	−2.004	0.18	−8.948**	−2.242	0.20	−8.354**	−1.684	0.19
	(3.680)	(3.729)		(3.644)	(3.713)		(3.549)	(3.588)	
Lottery	−8.181**	−5.117	0.56	−7.918**	−4.688	0.55	−6.977**	−3.548	0.54
	(3.551)	(3.842)		(3.516)	(3.778)		(3.496)	(3.655)	
Feedback (winner)	−10.482**	−4.965	0.49	−9.269**	−1.477	0.40	−9.965**	−2.516	0.41
	(4.320)	(6.651)		(4.327)	(6.492)		(4.255)	(7.105)	
Feedback (loser)	−13.561***	−6.234*	0.12	−14.728***	−7.649**	0.10	−14.615***	−6.950**	0.08
	(3.322)	(3.356)		(3.289)	(3.371)		(3.279)	(3.209)	
Performance				−0.308***	−0.493***		−0.263**	−0.390***	
				(0.108)	(0.117)		(0.104)	(0.125)	
Age				−0.103	−0.131		−0.061	−0.078	
				(0.123)	(0.102)		(0.120)	(0.098)	
Risk seeking							2.792***	3.442***	
							(0.449)	(0.463)	
R^2	0.018	0.006		0.027	0.024		0.069	0.094	
N	934	741		934	741		934	741	

Notes. Robust standard errors are in parentheses. Gender difference p-values are from regressions including gender interactions.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

of the effect of being a feedback loser, which is significantly larger for men than for women (p-values are from Wald tests after regressions that include gender–treatment interactions).

Past research has also shown that the performance of men reacts more strongly to competitive incentives than the performance of women. Although Figure 3 suggests that the treatments had no impact on performance, this could mask gender differences in the reaction to the competitive incentive schemes. This idea is explored in Table 7, where we regress performance in the slider task on treatment dummies separately for men and for women. The effects of the treatments on performance are not generally significant with the exception of the feedback treatment, which significantly lowers performance for women.[6] None of the effects differs significantly between men and women.

Potential Mechanisms

Our results show that exposing people to competitive or random payment schemes, as opposed to individual piece-rate pay, has a negative impact on their willingness to cooperate. The literature on behavioural spillovers suggests a number of potential mechanisms behind our results. One such candidate is cognitive load. Competition is a more complex situation compared to working under individual incentives because of added uncertainty and strategic concerns. These additional cognitive demands could lead to cognitive load whereby cognitive constraints might prevent individuals from choosing optimally (Samuelson 2001). Cognitive load might lead people to rely on simple heuristics such as playing Nash or choosing randomly, which in the case of the public goods game would lead to lower contributions. The literature on cognitive load in social preference games does not generally support this conjecture. Most relevant to our study, Duffy and Smith (2014) find a positive effect of cognitive load on cooperation levels in the prisoner's dilemma. If our treatments induce cognitive load, the results of Duffy and Smith would thus make us expect a positive effect on cooperation from our treatments, whereas we find the opposite.[7]

In Table 8, we further explore cognitive load as a potential mechanism by looking at whether individuals are more likely to make certain choices depending on the treatment. Here, we regress dummies for specific choices in the public goods game on treatment dummies and controls. Columns (1) and (2) show that participants in the three treatment groups are about five percentage points more likely to play Nash (i.e., to contribute zero) compared with the control group, but this effect is statistically significant only for the feedback treatment. Columns (3) and (4) show that

[6] A potential explanation for the negative impact of the feedback treatment on performance for women is fear of failure (Elliot and Thrash 2004). Individuals high on fear of failure seek to avoid failure in achievement settings, for example, by mentally escaping the threatening situation by withdrawing effort (Elliot and Church 1997). McGregor and Elliot (2005) find that women are more affected by fear of failure than men. Women have also been found to be more feedback averse than men (Möbius et al. 2011).

[7] Moreover, studies on giving in dictator games find either positive or zero effects (Cornelissen et al. 2011, Hauge et al. 2009, Schulz et al. 2012) and Cappelletti et al. (2011) find a zero effect on ultimatum game proposals.

Table 7 OLS Regressions by Gender (Dependent Variable: Performance in the Slider Task)

	(1) Male	(2) Female	*p*	(3) Male	(4) Female	*p*	(5) Male	(6) Female	*p*
Competition	0.656	−0.866	0.38	0.584	−0.577	0.48	0.520	−0.630	0.49
	(1.197)	(1.079)		(1.188)	(1.046)		(1.187)	(1.042)	
Lottery	1.132	0.578	0.77	1.005	0.799	0.91	0.904	0.670	0.91
	(1.163)	(1.114)		(1.155)	(1.079)		(1.154)	(1.076)	
Feedback	−1.007	−1.695*	0.63	−1.307	−1.806*	0.75	−1.290	−1.842**	0.73
	(1.041)	(0.961)		(1.037)	(0.931)		(1.035)	(0.927)	
Age				−0.150***	−0.201***		−0.154***	−0.205***	
				(0.039)	(0.028)		(0.039)	(0.028)	
Risk seeking							−0.286**	−0.363***	
							(0.144)	(0.133)	
R^2	0.006	0.009		0.021	0.072		0.026	0.081	
N	934	741		934	741		934	741	

Notes. Robust standard errors are in parentheses. Gender difference *p*-values are from regressions including gender interactions.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

treated participants are between 8 and 16 percentage points less likely to contribute the full endowment. This effect is significant for all three treatment groups. Columns (5) and (6) show that treated participants are slightly more likely to give exactly half their endowment, but again, this effect is significant only for the feedback group. In summary, the only treatment group for which we find strong evidence of a change in choosing focal amounts is the feedback group. Participants in that group are more likely to choose Nash and more likely to give half (but they are also less likely to give everything). However, as previous analysis shows, the difference between the feedback treatment and the other treatments mainly stems from the effect of being informed of a negative outcome. It is not obvious why losers should suffer from greater cognitive load than winners.

Table 8 Simple Heuristics in the Public Goods Game (OLS Regressions)

	(1) Zero	(2) Zero	(3) Maximum	(4) Maximum	(5) Half	(6) Half
Competition	0.056*	0.050	−0.083**	−0.076**	0.014	0.012
	(0.033)	(0.033)	(0.038)	(0.037)	(0.021)	(0.021)
Lottery	0.049	0.037	−0.092**	−0.080**	0.024	0.025
	(0.033)	(0.032)	(0.038)	(0.037)	(0.022)	(0.022)
Feedback	0.062**	0.064**	−0.162***	−0.166***	0.044**	0.047**
	(0.029)	(0.028)	(0.033)	(0.033)	(0.019)	(0.019)
Male		0.052**		0.027		−0.037**
		(0.022)		(0.025)		(0.016)
Performance		0.004***		−0.003***		0
		(0.001)		(0.001)		(0.001)
Age		0.002**		0		0.001
		(0.001)		(0.001)		(0.001)
Risk seeking		−0.026***		0.043***		−0.006**
		(0.004)		(0.004)		(0.003)
R^2	0.003	0.040	0.015	0.072	0.007	0.027
N	1,675	1,675	1,675	1,675	1,675	1,675

Notes. The dependent variable in columns (1) and (2) is a binary variable indicating a zero contribution in the public goods game. The dependent variable in columns (3) and (4) is a binary variable indicating a maximal contribution of 80 in the public goods game. The dependent variable in columns (5) and (6) is a binary variable indicating either a zero contribution or a maximal contribution of 80 in the public goods game. Robust standard errors are in parentheses.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

Cognitive load could also lead people to choose randomly in the public goods game. If this is the case, we would expect a higher variance in the allocations in the three treatments compared to the control treatment. We do not find evidence for this. A robust test for equality of variances across the four treatment groups detects no significant difference ($p = 0.335$), and the treatment with the lowest variance is the feedback treatment. We interpret these findings to indicate that participants' resorting to heuristics as a result of cognitive load is unlikely to be the mechanism behind our findings.

A related concept that is potentially relevant to our results is outcome entropy. Bednar et al. (2012), in an interesting study where participants play two simple 2×2 simultaneous move games at the same time, find significant spillover effects. For example, people who play a prisoner's dilemma at the same time as a game of chicken choose the off-diagonal options more often than people who play a prisoner's dilemma in isolation. They find that outcome entropy (the amount of variation in the choices), which they interpret as a proxy for the complexity and the cognitive load induced by a game, predicts the direction of the spillover: low-entropy games influence choices in high-entropy games. In our case the direction of the spillover effect is clear as the games are played in sequence. However, it is interesting to see whether outcome entropy (and therefore presumably cognitive load) varies across treatments. A robust test for equality of variances across the four treatment groups indeed finds that the variance in the score achieved in the slider task varies significantly across treatments ($p < 0.001$). However, there is no consistent pattern: the standard deviation of performance is similar in the control and competition treatments (11.7 and 11.4, respectively), highest in the lottery treatment (13.3), and lowest in the feedback treatment (8.7).

A second plausible mechanism to explain our results is priming.[8] Mental representations of a phenomenon can have an effect on behaviour outside the context of that phenomenon because limited cognitive abilities prevent people from always accessing the most relevant mental representations required for a task or decision (as discussed in, e.g., Al-Ubaydli et al. 2013). Mental representations that have been recently or regularly accessed can thus have an effect on behaviour and choices even if they are not relevant. Competing with others or receiving unequal payoffs could prime individuals with competitive or selfish norms. Moreover, those participants who work under comparative incentives are part of a group, which by itself might trigger norms of competitive behaviour.

To test this mechanism, we conducted an additional experiment on MTurk.[9] The design again consisted of two rounds, one of which was randomly paid out. In the first round, we randomly primed participants with competition-related concepts without actually making them compete, using the scrambled sentences paradigm (Srull and Wyer 1979). The second-round public goods game was identical.

In the first round, all participants were paid a piece rate for their performance in a sentence-unscrambling task where they had to form correct four-word sentences out of five randomly ordered words. They earned 10 cents for each of the 10 sentences they unscrambled correctly. Participants in the control group received 10 neutral sentences; for the treatment group, 5 of 10 sentences were replaced by sentences containing words such as "competition," "contest," "rivalry," "tournament," or "conflict" (see the online appendix, available as supplemental material at http://dx.doi.org/10.1287/mnsc.2015.2257, for a full list of sentences).

A further difference between comparative and individual incentives is the fact that workers are part of a group, which might in itself lead them to compare themselves with others and trigger competitive norms. We therefore randomly assigned half the participants in each the treatment and control groups to

[8] Chen et al. (2014) show that cooperation can be susceptible to priming by emphasizing different aspects of participants' identities in a lab experiment.

[9] Note that participants in the main experiment were excluded from participating in the additional experiment.

Table 9 **Descriptive Statistics for Supplementary Priming Experiment**

	Sample	Men	Women
Observations			
No prime	402	257	145
Competition prime	397	241	164
Group prime	405	231	166
Competition prime × Group prime	401	246	155
Total	1,605	975	630
Variable means			
Performance (number of unscrambled sentences)	9.3	9.2	9.4
PG allocation	40.2	38.0	43.5
Age	30.9	29.3	33.3
Risk seeking	6.4	6.7	6.1

Note. PG, public goods.

a group prime whereby the instructions informed the participants that "during the whole study, you will be in a group of four participants." This amounts to a 2×2 design (competition prime × group prime).

Table 9 shows descriptive statistics. We recruited an additional 1,605 participants on Mechanical Turk, referred them to an external website (Qualtrics), and randomly allocated them to one of the four treatments. The instructions and the questionnaire were as close as possible to the ones used in the previous experiment (see the online appendix). The average earnings in the sentence unscrambling task across all treatments were \$0.92, and the average earnings for the public goods game were \$1.17. Participants also received a fixed 50-cent fee for completing the experiment. The median amount of time it took participants to complete the experiment was six minutes. Participation in the experiment was restricted to individuals living in the United States.

Table 10 shows the results. Participants primed with competition phrases contribute approximately seven cents less to the public good than members of the control groups who unscrambled neutral sentences. Participants who received instructions containing the group prime also gave significantly less, an effect of approximately eight cents. Participants who were treated with both the group prime in the instructions and the competition prime in the unscrambling task gave approximately eight cents less and are therefore indistinguishable from the participants who only received one of the two primes. Contrary to the main experiment, the effects of gender and age are significant, whereas the effect of performance is insignificant.

These results support the interpretation that the negative effect of comparative payment schemes on the willingness to cooperate with others works through priming. The results of the additional experiment leave open whether the priming effect is due to the presence of others making people think more

Table 10 **Results from Supplementary Priming Experiment (OLS Regressions; Dependent Variable: Public Goods Game Contribution)**

	(1)	(2)	(3)
Competition prime	−6.929***	−7.267***	−6.338***
	(2.521)	(2.528)	(2.435)
Group prime	−7.920***	−8.005***	−8.309***
	(2.496)	(2.488)	(2.402)
Competition × Group prime	−7.965***	−8.022***	−6.955***
	(2.518)	(2.535)	(2.451)
Male		−4.866***	−6.847***
		(1.831)	(1.755)
Performance		−0.357	0.167
		(0.610)	(0.574)
Age		0.219***	0.260***
		(0.084)	(0.083)
Risk seeking			3.821***
			(0.319)
R^2	0.009	0.019	0.095
N	1,605	1,605	1,605

Note. Robust standard errors are in parentheses.
*, **, and *** denote significance at 10%, 5%, and 1%, respectively.

strategically or leading people to compare themselves to others. The results from the main experiment lend some support to the second explanation: providing feedback to participants significantly increases the treatment effect. Receiving feedback basically forces people who lost to compare their own performance and earnings with those of the winner (and vice versa for winners, for whom we also find a negative effect on cooperation compared with the control). It is also possible, however, that the feedback treatment strengthens a group priming effect.[10]

It is somewhat puzzling in general that the group prime decreases cooperation to the same extent as competition priming. This result suggests that it is the fact of being in a group with others rather than payment uncertainty that triggers the negative effect of comparative payment schemes on cooperation. A unifying explanation is that being in a group with others who earn different amounts leads people to compare themselves, which possibly triggers competitive or selfish norms that carry over to other settings. The results also support a slightly different explanation, however. Only in the risk treatment do the instructions explicitly mention to participants that "you are in a group with 3 other randomly selected participants. These are NOT the same individuals you interact with in the allocation part." As the competition priming and the simple group prime are both sufficient to recreate the effect, one could speculate that participants in the competition treatment give less because of competitive norms caused by competition priming and participants in the risk treatment give less because of the group prime. Although the additional experiment supports priming as the likely mechanism behind our results, the exact pathways are consequently still open for debate and thus more work.

Conclusion

In this study, we demonstrate that comparative payment schemes have a detrimental effect on the willingness of individuals to cooperate with others even when cooperation does not affect the chance of winning. Moreover, our design excludes reciprocity, beliefs, and income effects as explanations. We therefore interpret the measured impact of comparative pay on cooperation as a psychological spillover effect that exists over and above the potentially perverse incentive effects of comparative payment schemes. In a supplementary experiment we show that priming is a plausible psychological mechanism behind our results. That is, comparative payment schemes prime people with competitive norms, possibly by inciting them to compare themselves with others.

If our results generalise to a workplace context, they would have important ramifications for the design of optimal organisational structures and incentive schemes. Hierarchical organisational structures are inherently competitive because only a few of those on a given level will manage to move up to the next level. If the hierarchical pyramid is steep, employees on all levels find themselves in constant competition. In organisations that put a premium on cooperation, flat hierarchies are therefore more attractive. Likewise, comparative compensation schemes lead to employees constantly being forced to compare themselves with and compete against their coworkers. In organisations where cooperation is important, fixed wages, individual incentives, or team-based bonuses might be more appropriate.

In our experimental task, we find no effect of comparative reward schemes on performance. However, our task is very simple and the time frame is short, which might limit the scope for an effect of incentive schemes on performance. Economic theory, on the other hand, indicates that competitive incentives boost performance and that is why companies use them. Our results show that any potential productivity gain engendered by comparative reward schemes might come at the cost of a reduced willingness to cooperate. Which effect dominates will depend on the magnitude of the productivity gain and the importance of cooperation for the aims of the organisation. One can imagine that a company would not be unduly worried about the cooperativeness of, e.g.,

[10] One referee suggested that one could use our setup with a dictator game to disentangle this further, and we agree that this is interesting for future work.

independently operating salespeople. But when teamwork and organisation-wide cooperation are crucial for an organisation's goals, competitive incentives and promotion mechanisms could carry a huge cost.

Future work should explore more complex skill-based tasks where the effect of incentive schemes on performance might be different. To examine the robustness and external relevance of our results, replications using other forms of contests and other social preference games would be welcome. Furthermore, it would be interesting to know whether inequality in payoffs is enough in itself to trigger a negative effect on the willingness to cooperate. Finally, a field experiment might be most appropriate to examine whether spillovers that exist in the short term in the lab carry over to the long-term interactions typical in most organisations.

If one is willing to extrapolate even further, our results could be relevant for predicting how societies change as they become more competitive. Arguably, the lives of many people in advanced economies have recently become more competitive as a result of higher unemployment and lower job security. Our results could be relevant for our understanding of the recent financial crisis, which exposed many examples of extreme selfish behaviour in competitive environments. Our results suggest that this might not be due purely to a selection effect (competitive environments attracting selfish people) but might be partially explained by an effect of the work environment on the mind-set of the employees.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2015.2257.

Acknowledgments

The authors thank Gordon Kraft-Todd and Janna Ter Meer for help conducting the study on MTurk, and seminar participants at the European Meeting of the Econometric Society Toulouse 2014, Goethe University Frankfurt, the Swedish Economics Conference in Umeå 2014, University of Amsterdam, University of Aarhus, and the University of Zürich Norms and Cooperation Workshop; as well as Anne Boschini, Karin Hederos Eriksson, Magnus Johannesson, Johanna Möllerström, Astri Muren, Hessel Oosterbeek, Eva Ranehill, Randolph Sloof, and Mirjam van Praag for great comments. The authors also thank the Knut and Alice Wallenberg Foundation (A. Dreber is a Wallenberg Academy Fellow), the Jan Wallander and Tom Hedelius Foundation (Handelsbankens Forskningsstiftelser [Grant P2010-0133:1]), and the Speerpunt Behavioural Economics of the University of Amsterdam for financial support.

References

Al-Ubaydli O, Houser D, Nye J, Paganelli MP, Pan XS (2013) The causal effect of market priming on trust: An experimental investigation using randomized control. *PloS One* 8(3):Article e55968.

Amir O, Rand DG, Gal YK (2012) Economic games on the Internet: The effect of $1 stakes. *PloS One* 7(2):Article e31461.

Andrade EB, Ariely D (2009) The enduring impact of transient emotions on decision making. *Organ. Behav. Human Decision Processes* 109(1):1–8.

Balafoutas L, Sutter M (2012) Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science* 335(6068):579–582.

Bandiera O, Barankay I, Rasul I (2013) Team incentives: Evidence from a firm level experiment. *J. Eur. Econom. Assoc.* 11(5): 1079–1114.

Bednar J, Chen Y, Liu TX, Page S (2012) Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games Econom. Behav.* 74(1):12–31.

Brandts J, Riedl A, Winden FV (2009) Competitive rivalry, social disposition, and subjective well-being: An experiment. *J. Public Econom.* 93(11):1158–1167.

Buser T (2014) The impact of losing in a competition on the willingness to seek further challenges. Discussion Paper 14-083/I, Tinbergen Institute, Amsterdam.

Cappelen AW, Sørensen EØ, Tungodden B (2013) When do we lie? *J. Econom. Behav. Organ.* 93(September):258–265.

Cappelletti D, Güth W, Ploner M (2011) Being of two minds: Ultimatum offers under cognitive constraints. *J. Econom. Psych.* 32(6):940–950.

Cason TN, Gangadharan L (2013) Cooperation spillovers and price competition in experimental markets. *Econom. Inquiry* 51(3):1715–1730.

Charness G, Masclet D, Villeval MC (2013) The dark side of competition for status. *Management Sci.* 60(1):38–55.

Chen K-P (2003) Sabotage in promotion tournaments. *J. Law, Econom., Organ.* 19(1):119–140.

Chen DL (2010) Markets and morality: How does competition affect moral judgment? Working paper, Duke University School of Law, Duke University, Durham, NC.

Chen Y, Li SX, Liu TX, Shih M (2014) Which hat to wear? Impact of natural identities on coordination and cooperation. *Games Econom. Behav.* 84(March):58–86.

Cornelissen G, Dewitte S, Warlop L (2011) Are social value orientations expressed automatically? Decision making in the dictator game. *Personality Soc. Psych. Bull.* 37(8):1080–1090.

Croson R, Gneezy U (2009) Gender differences in preferences. *J. Econom. Literature* 47(2):448–474.

Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG (2011) Individual risk attitudes: Measurement, determinants, and behavioral consequences. *J. Eur. Econom. Assoc.* 9(3): 522–550.

Duffy S, Smith J (2014) Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *J. Behav. Experiment. Econom.* 51(December):47–56.

Elliot AJ, Church MA (1997) A hierarchical model of approach and avoidance achievement motivation. *J. Personality Soc. Psych.* 72(1):218–232.

Elliot AJ, Thrash TM (2004) The intergenerational transmission of fear of failure. *Personality Soc. Psych. Bull.* 30(8):957–971.

Falk A, Szech N (2013) Morals and markets. *Science* 340(6133): 707–711.

Falk A, Fehr E, Huffman D (2008) The power and limits of tournament incentives. Working paper, University of Bonn, Bonn, Germany.

Gill D, Prowse V (2012) A structural analysis of disappointment aversion in a real effort competition. *Amer. Econom. Rev.* 102(1):469–503.

Gill D, Prowse V, Vlassopoulos M (2013) Cheating in the workplace: An experimental study of the impact of bonuses and productivity. *J. Econom. Behav. Organ.* 96(December):120–134.

Gneezy U, Rustichini A (2005) Executive versus teachers: Gender, competition and self-selection. Working paper, University of California, San Diego, La Jolla.

Gneezy U, Niederle M, Rustichini A (2003) Performance in competitive environments: Gender differences. *Quart. J. Econom.* 118(3):1049–1074.

Goette L, Huffman D, Meier S, Sutter M (2012) Competition between organizational groups: Its impact on altruistic and antisocial motivations. *Management Sci.* 58(5):948–960.

Harbring C, Irlenbusch B (2011) Sabotage in tournaments: Evidence from a laboratory experiment. *Management Sci.* 57(4): 611–627.

Harlé KM, Sanfey AG (2007) Incidental sadness biases social economic decisions in the ultimatum game. *Emotion* 7(4): 876–881.

Hauge KE, Brekke KA, Johansson L-O, Johansson-Stenman O, Svedsäter H (2009) Are social preferences skin deep? Dictators under cognitive load. Working Paper in Economics 371, University of Gothenburg, Göteborg, Sweden.

Herz H, Taubinsky D (2013) Market experience as a reference point in judgments of fairness. Working paper, University of Fribourg, Fribourg, Sweden.

Horton JJ, Rand DG, Zeckhauser RJ (2011) The online laboratory: Conducting experiments in a real labor market. *Experiment. Econom.* 14(3):399–425.

Konrad KA (2000) Sabotage in rent-seeking contests. *J. Law, Econom., Organ.* 16(1):155–165.

Lazear EP (1989) Pay equality and industrial politics. *J. Political Econom.* 97(3):561–580.

McGregor HA, Elliot AJ (2005) The shame of failure: Examining the link between fear of failure and shame. *Personality Soc. Psych. Bull.* 31(2):218–231.

Möbius MM, Niederle M, Niehaus P, Rosenblat TS (2011) Managing self-confidence: Theory and experimental evidence. NBER Working Paper 17014, National Bureau of Economic Research, Cambridge, MA.

Mollerstrom J (2013) Quotas and cooperation. Working paper, Harvard University, Cambridge, MA.

Niederle M, Vesterlund L (2007) Do women shy away from competition? Do men compete too much? *Quart. J. Econom.* 122(3): 1067–1101.

Samuelson L (2001) Analogies, adaptation, and anomalies. *J. Econom. Theory* 97(2):320–366.

Savikhin AC, Sheremeta RM (2013) Simultaneous decision-making in competitive and cooperative environments. *Econom. Inquiry* 51(2):1311–1323.

Schulz JF, Fischbacher U, Thöni C, Utikal V (2012) Affect and fairness: Dictator games under cognitive load. *J. Econom. Psych.* 41:77–87.

Srull TK, Wyer RS (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *J. Personality Soc. Psych.* 37(10):1660–1672.

Ter Meer J (2014) The indirect effect of monetary incentives on deception. CGS Working Paper 05-04, University of Cologne, Cologne, Germany.