Better for Less DISCUSSION PAPER SERIES

DP No. 02/2016

(updated November 2017)

The Impact of Relative Performance Feedback on Beliefs and Preferences across Dissimilar Tasks

Lingbo Huang

Department of Economics, Monash University, Wellington Road, Clayton VIC 3800, Australia
lingbo.huang@outlook.com

Zahra Murad

School of Economics, University of Surrey, GU2 7XH, United Kingdom

z.murad@surrey.ac.uk

School of Economics, Faculty of Arts and Social Sciences, University of Surrey, Elizabeth Fry Building 04 AD 00, Guildford, GU2 7XH, Surrey, UK. T: +44(0)01483686623, E-mail: deliveringbetter@surrey.ac.uk, Web: www.deliveringbetter.com.

The Impact of Relative Performance Feedback on Beliefs and Preferences across Dissimilar Tasks

Lingbo Huang, Zahra Murad*

Abstract

Employees typically work on multiple tasks that require unrelated skills and abilities. While past research strongly supports that feedback influences beliefs and preferences for comparative pay within same tasks, little is known about effects of feedback across different tasks. In a novel laboratory experiment using two unrelated real effort tasks, we find that feedback of relative performance in the first task significantly affects valuations of (i.e. revealed preferences for) comparative pay in the second task. Further analysis suggests that a taste-based mechanism and a belief-based mechanism play independent roles in explaining the observed feedback effects. We also find that relative performance feedback on average increases females' valuations of comparative pay compared to when such feedback is absent. The results have important implications for organizations to understand both the powers and the limitations of using relative performance feedback as intervention policies.

Keywords Relative Performance Feedback, Confidence, Comparative Pay, Behavioral Spillovers, Gender

JEL classification C72 C91

L. Huang Department of Economics, Monash University, Wellington Road, Clayton VIC 3800, Australia. e-mail: lingbo.huang@outlook.com
* Corresponding Author: Z. Murad School of Economics, University of Surrey, Guilford, GU2 7XH, United Kingdom. e-mail: zahra.murad55@gmail.com

Acknowledgements We thank seminar participants at 7th Southern Experimental Economists Meeting in Malta, ASFEE 2016 meeting in France, Jo Blanden, Nigel Burnell, Leonie Gerhards, Aidas Masiliūnas, Peter Norman Sorensen, and Roel van Veldhuizen for valuable comments and suggestions. We acknowledge the financial support from the Leverhulme Trust (Grant RL-2012-681).

1 Introduction

Modern workplaces increasingly require employees to multitask across an array of tasks and job responsibilities that involve different skill sets (Coviello, Ichino, & Persico, 2014; Schottner, 2008). For example, academics often have multiple job responsibilities such as teaching, researching, leading teams and administration processes that have been argued to require different skills and abilities (Gautier & Wauthy, 2007). Other examples include, employers in consultancy and high-tech industries often using job rotation schedules to improve employees' motivation and to decrease burn-out (Arya & Mittendorf, 2004; Hsieh & Chao, 2004); and sales representatives employing different sales strategies to promote a variety of products and services to customers with different needs and personalities (Johlke, 2006). In such workplaces, relative performance feedback received in one task, which is used for the current performance evaluation, may have unintended impacts on confidence, mood, persistence, goal pursuit, effort and risk-taking in subsequent unrelated tasks. This can create path dependence such that outcomes achieved in one task lead to exaggerated long-term consequences in other tasks. In the light of the prevalence of such multitasking situations, we investigate in a laboratory experiment whether relative performance feedback in the first work task spills over and affects confidence (beliefs) and valuations of (revealed preferences for) comparative pay in the second unrelated work task. And if so, what the mechanism(s) are underlying the observed feedback spillover effects.

The existing experimental literature on relative performance evaluations and feedback has thus far overwhelmingly focused on feedback effects within the same task (e.g., Gill, Kissova, Lee, & Prowse, 2016; Kuhnen & Tymula, 2012; Möbius, et al, 2014; Buser, Gerhards, & van der Weele, 2017). The central novelty of our study is assessing the effects of relative performance feedback using two substantially different real effort tasks while keeping the work environments comparable. This allows us to study the exogenous effects of feedback by disentangling between the informational and motivational effects of feedback on valuations of comparative pay. We hypothesize two mechanisms to operationalize the feedback effects: a belief-based mechanism predicts that feedback affects individuals' beliefs about their (general) task abilities relative to their group members which in turn affect their valuations of comparative pay. A taste-based mechanism however suggests that feedback directly affects individuals' tastes for social comparisons that are inherent in comparative pay, independent of any feedback effects on their beliefs. To our knowledge this is the first paper to disentangle between the two mechanisms underlining the effects of relative performance feedback and thus contributes to the extensive literature on feedback effects.

We test our hypotheses in a three-part laboratory experiment, throughout which participants are matched in fixed groups of four. Participants work on a visual perception task in the first part and then on a mathematical task in the second part. In a between-subject design, we manipulate whether participants receive relative performance feedback within their group after the first task. In both parts, we reward participants' performance by piece rate and elicit their confidence relative to other group members. In the third part of the experiment, we elicit participants' valuations for submitting their second task performance to a comparative pay scheme.

The design choice of measuring valuations of comparative pay was motivated by the observation that most work decisions such as job, award and promotion applications are taken given individuals' previous performance history (Boring & Brown, 2016; Pohlhaus, et al, 2011). Such decisions are usually taken on the basis of an individual's previous performance where the selection procedure is based on relative evaluations of *past* performance and does not involve a subsequent competitive task. In our view, valuations of comparative pay serve as an intermediary between confidence as a belief and a preference to perform a task under comparative pay. We believe, such preferences have been understudied in economics and business literature and our study aims to fill this gap.

While studying the effect of feedback on valuations of comparative pay, we particularly focus on gender. Previous studies suggest that performance feedback as an intervention policy helps to mitigate gender differences in confidence and willingness to compete (Ertac & Szentes, 2011; Wozniak, Harbaugh, & Mayr, 2014, 2015). Establishing how male and female employees respond to feedback across multiple tasks is crucial to understanding how organizations should design their feedback policies given the gender composition of their workforce. The choice of the specific tasks that we use in our experiment, detailed in Section 3, is also partially motivated by previous experiments on gender differences in confidence and preferences for comparative pay.

Section 4 presents the experimental results. We observe significant effects of relative performance feedback from the first task on valuations of comparative pay in the second task. Individuals who receive positive feedback in the first task show significantly higher valuations for submitting second task performance to comparative pay than those who receive negative feedback. We also detect some evidence of feedback effects on confidence. Importantly, we identify that the observed effect of feedback on valuations operates independently of feedback effects on confidence beliefs, thus supporting the belief-based and the taste-based mechanism play important and independent roles in explaining our data. We also find that the feedback

effects have gender components. Consistent with previous findings, when relative performance feedback is unavailable, women show lower valuations of comparative pay than men. The gender difference decreases (or in some cases reverses) when subjects receive relative performance feedback.

Our research differs from the literature by focusing on the effects of feedback across two different tasks, rendering the feedback uninformative of future performance and thus allowing us to investigate the *motivational* effects of feedback on confidence and valuations of comparative pay. We are interested in identifying whether the observed feedback effect is taste-based or belief-based. Buser (2016) also studies feedback effects on behavior in a subsequent task; but instead of studying feedback effects across two different tasks, he considers feedback effects across two environments with different incentive structures. Specifically, he investigates whether wins or losses in a previous competition have impacts on goal-setting behavior and actual performance in a subsequent (same) task under piece rate incentives with self-imposed target thresholds.[1]

Our paper also links to the literature on behavioral spillovers across strategic games and individual decision-making environments. Laboratory experiments have documented that prior experience in one game can spillover, resulting in a more cooperative and coordinated behavior in related games where cooperation and coordination is usually not observed (Ahn, et al., 2001; Cason et al., 2012; Knez & Camerer, 2000). In individual decision-making environments, Dolan & Galizzi's (2015) review of literature shows that behavioral spillovers can occur when an intervention targeting one individual or one type of behavior, spills over to affect other individuals or other types of behavior. For example, Bruhin, et al., (2015) show that an intervention in the form of a phone call to increase blood donations of individuals spills over to increase donation rates of those in close proximity of the phone call receivers. In our setup, we define a "feedback spillover" as having occurred when an information pertaining to one environment spills over to affect beliefs and preferences pertaining to another sufficiently different environment. While we benefit from adopting some of the design features used in previous studies on behavioral spillovers, our study is the first to manipulate feedback and test for its spillover effects across two individual decision-making environments.

The individual decision making environment that we focus on, with some manipulations, can be extended to dynamic tournaments, as shown by Ederer (2010). Hence, our study can also be contrasted with the literature on feedback effects in dynamic contexts, for example, promotion

[1] See Fishbach, Eyal, & Finkelstein (2010) for a review of feedback effects on goal pursuit.

tournaments (Brown & Minor, 2014) and political campaigns and elections (Klumpp & Polborn, 2006; Malueg & Yates, 2010). In those contexts, positive feedback of past performance endogenously predicts a higher probability of winning in future tournaments against the same opponents, thus building up a success or failure path that depends on the initial outcome. By changing the tasks across the two parts, we eliminate any strategic or endogenous effects of feedback. Furthermore, by fixing the group matching across the three parts of the experiment, we strengthen the psychological inter-correlation between the parts.

The results of this study have important implications on the social desirability of feedback policies that employers may utilize to promote more efficient multi-task workplaces. While in some situations, feedback may be a useful tool to eliminate biases and inform agents of their abilities within the workplace, if not implemented properly it may also create unintended and undesirable spillover effects onto beliefs and preferences in other tasks. We discuss the results and conclude in Section 5.

2 Experimental Design

To test our hypotheses, we ran two sets of pilot experiments and two sets of main experiments in the course of two academic years at the University of Surrey. In total, 42 subjects participated in the pilot experiments (male = 50%, age: M=21.5, SD=4.63) and 320 subjects participated in the main studies (male=50%; age: M=21.9, SD=5.85). The experiments with embedded instructions were programmed in z-Tree (Fischbacher, 2007).[2] A typical session lasted around forty-five minutes with the average payment of £13.70, including a £2 show-up fee and £5 for completing the experiment. Study 2 was conducted one year after Study 1 was completed and served as a further robustness check for replicability of Study 1 results with slightly modified experimental design features.

2.1 Pilot Experiments

The validity of our experimental design depends on a sufficient dissimilarity between two work tasks that are ideally unrelated in every aspect. For the second task, we chose to use a number adding task consisting of adding up sets of five randomly generated two-digit numbers. Subjects had five minutes to complete as many summations as possible. Calculators were not allowed but subjects could use provided scratch paper. After submission of each answer, a new problem was

[2] All instructions and the post-experimental questionnaire are reproduced in the Appendix A.

shown without any feedback on whether the former answer was right or wrong. Right answers added up to subject's final performance score and wrong answers were not penalized.

We chose number adding task as our second task because previous research suggests that while women are as competent as men in number adding tasks, they are less confident and less willing to enter competitions when their performance is judged relative to others (Ifcher & Zarghamee, 2016; Kamas & Preston, 2012; Niederle & Vesterlund, 2007). One of our interests in studying feedback spillover effects is to assess whether and how gender differences in confidence and valuations of comparative pay react to feedback exogenously. The gender neutrality in performance but not in beliefs and preferences enabled us to study feedback effects on gender differences independent of actual performance.

For the first task, in order to make an empirically grounded selection of unrelated tasks to the number adding task, in Pilot 1 we performed a pre-test evaluation of 5 different experimental tasks. These were the circle task (Hollard, Massoni, & Vergnaud, 2016), the counting zeros task (Abeler, et al, 2011), the ball-catching task (Gächter, Huang, & Sefton, 2016), the slider task (Gill & Prowse, 2012) and the number adding task (single-digit variant). Pilot 1 involved 18 subjects who were asked to complete all five tasks presented to them in random order. They were paid according to a piece rate in one randomly selected task.[3] For each task, we measured subjects' performance, their perceptions of task difficulty on a 7-point scale, and their choices of submitting past performance to comparative pay. Pilot 2 involved 24 subjects who were asked to complete the circle task, the number adding task (double-digit variant), the counting zeros task, the slider task and the anagram task (Buser, Gerhards, & van der Weele, 2017). In Pilot 2, we also asked subjects to rate the importance of five cognitive skills for each task performance and whether they thought that the performance between combinations of two tasks are related. We asked for the importance of Attention (holding attention while completing the task), Working Memory (ability to learn information and use that information for the current activity), Visual Perception (ability to see and interpret the visual information), Cognitive Flexibility (being able to consider several solutions or plans, not only the first one that comes to mind) and Numeracy Skills (Ability to reason and to apply simple numerical concepts) for each of the five tasks subjects completed. The instructions for the pilot experiments are presented in the Appendix A2; the results of the two pilot experiments are presented in the Appendix B. Based on a within-subject analysis of the pilot data (summarized in the Table B1), we decided to use the circle task as our first task: the correlation between the circle task and the number adding

[3] Piece rates were varied across the tasks to equalize expected earnings from each task.

task was not significant either in performance, perceptions of difficulty or choices of comparative pay. Most importantly, in the domains of visual perception and numeracy skills, the highest rated skills for success in the circle and the number adding task respectively, subjects rated the two tasks as unrelated. Furthermore, relative confidence and actual performance in the circle task were gender-neutral, ensuring that feedback received in the circle task would be independent of subjects' gender.[4]

It is worth noting that statistically insignificant correlation across tasks based on a finite sample will be subject to both type I and type II errors and it may not imply that the two tasks are necessarily unrelated. While this may be the case, in addition to the fact that we chose two substantially different tasks that require cognitive skills in different human faculties, we took an informed choice out of five other real-effort tasks that have previously been used in experiments studying confidence beliefs and valuations of comparative pay. The insignificant and a low correlation between the number adding and the circle tasks was later confirmed in our main experiments using a larger dataset from 320 subjects. We also elicited subjects' perception of the relatedness of the tasks in the main experiments which we use as controls together with the actual performance in both tasks when testing for feedback effects across the tasks.

The circle task is a visual perception task, which requires subjects to see a pair of black circles with white dots in them for 1 second and to judge which circle has more dots (see Figure 1). The version of the circle task used in our experiment consisted of 60 such pairs of circles. The position of dots in every pair of circles was randomly determined. But one circle always contained 50 dots, and the other 55 dots. Subjects were asked to choose the circle with more dots for every pair of circles. After submission of each answer, a new pair was shown without telling them whether the former answer was right or wrong. Correct answers added up to subject's final performance score; wrong answers were not penalized.[5]

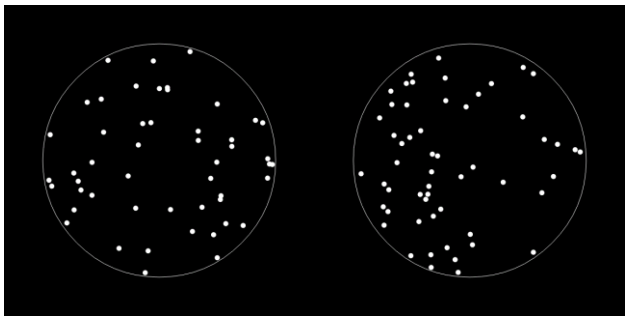
2.2 Main Studies

Study 1 was conducted in October 2015 with 168 subjects (male=54.7%; age: M=22.6, SD=6.07). Study 2 was conducted in June and October 2017 with 152 subjects (male=45.4%; age: M=21.1, SD=5.51). Upon arriving at the lab, subjects were seated according to randomly allocated ID numbers they received from the experimenter. The general instructions of the

[4] The slider task was another task that could be used, also uncorrelated with the number adding task in all of the three respects. But performance in the slider task is not gender-neutral (Gill & Prowse, 2014).

[5] It is worth noting that as the circles were shown to subjects for 1 second, it was impossible to even attempt to count the dots, hence there is no parallelism between the two tasks in terms of measuring counting ability. For a more in detail description of the task please refer to Fleming et al. (2016) and Hollard et al. (2016).

experiment, including general participation rules, rules for cash payment and consent forms, were provided in paper form and were read aloud by the experimenter. Subjects were told that they would be asked to complete several parts and that one part would be randomly selected for payment at the end of the experiment. They were also told that the specific instructions for different parts of the experiment would be shown on their computer screens before the beginning of each part.



Which circle has more dots?

Left Right

Figure 1: The circle task

At the beginning of a session, each subject was randomly matched into a group with three other subjects and the group remained the same throughout the session. In Part 1, all subjects worked on the first task—the circle task, and in Part 2 they worked on the second task—the number adding task. Performance in both tasks was incentivized by piece rates: £0.20 for each correctly judged circle and £0.50 for each correctly solved number addition. In a between-subject design, we manipulated whether subjects received feedback about their relative standing within their group according to their circle task performance at the end of Part 1. In the *NoFeedback* treatment of Study 1, subjects received feedback only on their absolute performance score in the circle task. In the *Feedback* treatment of Study 1, in addition to their absolute performance score, subjects also received feedback on their relative performance, which read as "your score was one of the TOP/BOTTOM two scores of your group." In Study 2, we chose not to give subjects in either treatment any information on absolute performance scores in the Circle task. By withholding absolute performance feedback after the circle task and only varying whether or not subjects received relative performance feedback on the first task, we removed a possible mitigating factor that might have prevented us from identifying significant effects of relative performance feedback across tasks. The potential mitigating factor was that if absolute performance feedback somehow helped subjects infer their relative performance, our manipulation of relative performance feedback might not be as salient as we have expected.

Before each task, subjects were asked to judge, on a scale from 0% to 100%, the probability that their performance score would be one of the top two scores of their group. Specifically, they were asked to assign 0% if they were completely certain about scoring in the bottom two of their group; assign 100% if they were completely certain about scoring in the top two of their group; and assign intermediary values to the degree they deemed appropriate for their uncertainty. They could indicate their choices using the slider on their screens, which could be moved from 0 to 100. The answer to this question served as the *pre-task confidence* measure. After they completed the number adding task and were informed of their actual performance score, they were asked the same question about whether their performance score was one of the top two scores in their group, the answer to which served as the *post-task confidence* measure. In Study 1, we did not incentivize the elicitation, while in Study 2 we incentivized the reported beliefs with Becker-DeGroot-Marschak (BDM) mechanism.

We used the two pre-task confidence measures to determine whether and how relative performance feedback modulated confidence in the second task, conditional on the confidence in the first task. By eliciting both pre- and post-task confidence in the second task, we determined how persistent the effect of feedback on confidence was and whether actual task performance affected confidence. Our confidence elicitation method was robust to Benoit & Dubra's (2011) and Benoit, Dubra, & Moore's (2015) critique of "apparent overconfidence." To obtain more informative measures of confidence, they recommended using (i) well-defined groups that subjects could compare themselves to, (ii) well defined performance tasks, and (iii) eliciting confidence as a subjective belief distribution rather than expected rankings within the group. Our confidence elicitation tool satisfied all three criteria.[6] Our tool was a straightforward one-item question in which there was no clear incentive to misreport.

In Part 3, subjects did not have to perform a task, but were asked to fill in the multiple-choice list table as shown in Figure 2. The table comprised ten rows, and the choice in each row was between a tournament (Option A) and a piece rate (Option B) compensation. The chosen compensation option would be applied to their piece rate performance in the number adding task of Part 2. The tournament option paid £1.00 per correct number addition if subjects were in the top half of their group and £0.00 if they were in the bottom half. The piece rate option had ten levels in a decreasing order from £1.00 to £0.10 per correct addition. Subjects had to make a

[6] Criteria (iii) was satisfied assuming that subjects had additive subjective probabilities on scoring in the top and bottom halves of their groups.

choice for every row of the table. At the end of the experiment, one of the rows was randomly selected and the choice in that row was implemented for payment.

The method is a finer-grained measure of valuations of comparative pay compared to Niederle & Vesterlund's (2007) simple binary choice method. Given the construction of the table, subjects were expected to choose piece rate in the first row because it was the highest and the most certain payment for any performance level. At some point, they were likely to switch from piece rate option to tournament option, and the level of piece rate where they made the switch was the *piece rate (PR) equivalent* of the comparative pay. For example, a subject switching from piece rate to tournament on the fifth row implied that he/she valued the comparative pay scheme equivalent to the piece rate worth between £0.50 and £0.60. For simplicity, we take the lower bound of the equivalence range in the data analysis.

Row	A: Tournament	Your Choice	B: Piece Rate
1	You get £1.00 per correctly solved number addition if your score in Part 2 was in the TOP two scores of your group, and £0.00 if your score in Part 2 was in the BOTTOM two scores of your group.	A □ □ B	£1.00 per correct number addition
2		A □ □ B	£0.90 per correct number addition
3		A □ □ B	£0.80 per correct number addition
4		A □ □ B	£0.70 per correct number addition
5		A □ □ B	£0.60 per correct number addition
6		A □ □ B	£0.50 per correct number addition
7		A □ □ B	£0.40 per correct number addition
8		A □ □ B	£0.30 per correct number addition
9		A □ □ B	£0.20 per correct number addition
10		A □ □ B	£0.10 per correct number addition

Figure 2: Eliciting valuations of comparative pay in Part 3

We did not ask subjects to complete a further work task after they had filled in the table in Part 3. Thus we only measured *retrospective* preferences for comparative pay, rather than the willingness to compete. This part was similar to Ifcher & Zarghamee (2015)'s Task 5 and Niederle & Vesterlund (2007)'s Task 4, in both of which the choice of compensation was applied to subjects' past piece rate performance. Applying comparative pay to past performance also

ensured that social preferences did not play a role in subjects' choices since subjects' choices and performance would not affect the earnings of any other subjects; nor would their earnings depend on other subjects' choices of compensation.

	General Instructions	General participant rules; processing of payment; consent form
Part 1	Pre-task Confidence Elicitation	On a scale of 0% to 100%, indicate the probability that your score will be in the top half of your group. Un-incentivized [Study 1] BDM Incentivized [Study 2]
	The Circle Task	Complete 60 pairs of circles; £0.20 per correct answer
	Feedback	"NoFeedback" [Study 1]: absolute performance score "No Feedback" [Study 2]: no information about performance "Feedback" [Study 1]: absolute performance score + relative performance (top/bottom half of your group) "Feedback" [Study 2]: relative performance (top/bottom half of your group)
Part 2	Pre-task Confidence Elicitation	On a scale of 0% to 100%, indicate the probability that your score will be in the top half of your group. Un-incentivized [Study 1] BDM Incentivized [Study 2]
	The Number Adding Task	Add up as many as possible five two-digit numbers in 5 minutes; £0.50 per correct answer
	Feedback	Absolute performance score of Number Adding
	Post-task Confidence Elicitation	On a scale of 0% to 100%, indicate the probability that your score was in the top half of your group. Un-incentivized [Study 1] BDM Incentivized [Study 2]
Part 3	Valuations of comparative pay	Choose between the tournament and piece rate compensation for each of the 10 levels of piece rate from £0.10 to £1.00 to be applied to Part 2 score; one of the 10 choices for payment
	Questionnaire	Demographics; general risk/confidence/competitiveness; perceived relatedness of the two tasks
	Final Feedback	Absolute performance score + relative performance (top/bottom) in your group for both tasks. Final earnings.

Figure 3: Design of the Main Studies

Similar to the confidence measure, our measure of valuations of comparative pay could be influenced by other factors such as feedback aversion, risk and competitive attitudes.[7] We control for feedback aversion by making it clear that although all subjects will be paid by piece rates in Part 1 and 2, they would still receive relative performance feedback at the end of the

[7] However, note that these two measures are not identical as long as subjects are not expected value maximizers. Under non-linear expected utility, rank-dependent utility, non-neutral ambiguity models or other utility formulations that explicitly capture valuations of comparative pay, the PR-equivalents are not equivalent to measuring beliefs. Consistent with Ifcher & Zarghamee (2015) who called their measure the "price" of a competition, we refer to the elicited PR-equivalents in our context as valuations of, or revealed preferences for, comparative pay.

experiment in both treatment conditions.[8] We also explicitly measured general risk, confidence and competitive attitudes in an end-of-study questionnaire and subsequently used them as controls in the analysis of preferences.

After Part 3, subjects were asked to complete a questionnaire on demographic characteristics, self-reported economic preferences (general risk, confidence, and competitive attitudes), and perceived relatedness between the circle task and the number adding task.[9] At the end of the experiment, all subjects received information about their absolute performance scores and whether the scores were in the top or in the bottom half of their groups in each of the two tasks. They were also informed which part was selected for payment (in the case of Part 3, which row was selected) and their final cash payment. Figure 3 summarizes the design of the main studies.

3 Hypotheses

The aim of this paper is to assess the spillover effects of relative performance feedback on confidence and preferences for comparative pay across different tasks. In this section, we discuss two potential mechanisms that can operationalize these effects and state corresponding hypotheses.

The first mechanism relies on a belief-based approach that most studies in the literature on relative performance feedback follow, albeit in the framework of feedback within the same task. For example, providing subjects with full feedback on their own and others' performance repeatedly over the course of the experiment has been shown to eliminate biases in confidence (Moore & Cain, 2007; Murad, 2016; Rose & Windschitl, 2008).[10] While our design deliberately selects two unrelated tasks in an attempt to eliminate such belief updating about *specific* ability in a task, it remains possible that feedback effects operate through updating beliefs (e.g., standard

[8] Past research show that feedback aversion affects confidence levels, risk preferences, task choices, and effort levels in a variety of decision situations (Burks et al., 2013; Niederle & Vesterlund, 2007; Zeelenberg, et al., 1996). Thus giving subjects feedback on their relative performance at the end of experiment, irrespective of their choices in Part 3, will remove any potential effects of feedback aversion on choices in Part 3.

[9] The self-reported economic preferences were similar to the validated survey instruments developed by Dohmen et al. (2011) who showed that self-reported measures of six economic preferences strongly correlates with many real life and laboratory decisions.

[10] Feedback about subjects' rankings in the performance distribution (top, middle or bottom) can also cause asymmetric belief updating to overweight positive feedback and underweight negative feedback (Ertac, 2011). Similar asymmetric belief updating is also evident when feedback comes from bilateral comparisons with another randomly selected participant (Möbius, et al., 2014). In Eberlein, Ludwig, & Nafziger's (2011) study of heterogeneity in subjects' reactions to feedback, some subjects tend to disregard feedback while others overreact to it, rendering feedback ineffective in correcting overall biased beliefs. These asymmetric processes of belief updating have been argued to be motivated by the tendency to preserve ego or self-esteem (Kőszegi, 2006; Kuhnen & Tymula, 2012; Möbius et al., 2014).

Bayesian or ego-biased) of *general* ability relative to the reference group: *I am better than my peers in my group in a visual perception task → I am more able than my peers → I am better than my peers in my group in the numeracy task → I submit my performance to comparative pay*.[11]

The second mechanism adopts a taste-based approach in which feedback interacts with the taste for social comparisons. We hypothesize that positive feedback increases individuals' tastes for comparisons while negative feedback decreases tastes for comparisons. In our terminology, the taste for comparisons can reflect an individual's tendency of seeking or avoiding further competitions or interpersonal comparisons, representing either switching to comparative pay or sticking to a piece rate respectively in our experiment. Differently from the *cognitive* belief based mechanism taste based mechanism is *affective* in nature: *I am told I was better than my peers in my group in a visual perception task → I enjoyed how it felt to receive positive feedback → I will seek more comparison with my peers in the numeracy task → I submit my performance to comparative pay*. The claim that (unobserved) tastes for comparisons are malleable to the environment has found some support in experiments that examine explicit behavior of seeking or avoiding information on social comparisons by under- or over-confident individuals. For instance, Burks et al. (2013) find that individuals who are confident in their ability are more interested in learning about their relative performance. In our experimental framework, feedback may thus operate on its own, independent of its effect on confidence: subjects receiving negative feedback want to avoid comparisons of performance with their reference group and thus have lower valuations of comparative pay. Conversely, subjects receiving positive feedback want to seek comparisons and thus have higher valuations of comparative pay.

In sum, these two mechanisms generate the following distinct hypotheses:

Belief-based hypothesis: *Relative performance feedback in the first task will affect valuations of comparative pay in the second task through affecting relative confidence.*

Taste-based hypothesis: *Relative performance feedback in the first task will affect valuations of comparative pay in the second task independent of its effects on relative confidence.*

[11] A closely related line of experimental research about feedback effects on willingness to compete sheds light on the last logical induction. Kamas & Preston (2012), Dreber, von Essen, & Ranehill (2014) and Veldhuizen (2016) identify confidence of expected relative performance to be the most important determinant in entry decisions to competitions, and conditional on task performance, confidence beliefs account for most of the variations in gender gaps in willingness to compete.

By simultaneously analyzing feedback effects across tasks on confidence and valuations of comparative pay, we can determine which mechanism is the more plausible description of our data.

4 Results

The final sample for data analysis includes 168 subjects in Study 1 (NoFeedback: N=60; Feedback: N=108) and 152 subjects in Study 2 (NoFeedback: N=52; Feedback: N=100).[12] Since we do not impose a single switch point in Part 3, 4 subjects in Study 1 and 15 subjects in Study 2 who switched non-monotonically are removed from our analysis of PR-equivalents.

Before the main analysis, we check whether, in line with our pilot results, scores in the two tasks are indeed uncorrelated. The Pearson's correlation coefficients between the circle task and the number adding task scores are 0.080 (p=0.155), thereby reassuring the internal validity of our design. We also do not observe any correlation between being in the top half of the group in the circle and number adding tasks: Pearson's correlations coefficient between these two measures is -0.013 (p=0.824). Hence, there is no reason to believe that subjects who received positive or negative feedback in the circle task were more likely to be in the top or bottom halves of their groups in the number adding task. In other words, the feedback received before subjects begin to work on the number adding task can be seen as good as random.

4.1 Evidence on feedback spillover effects

Figure 4 presents our results of the three outcome variables for all feedback conditions of the two studies. The pre-task confidence levels which subjects reported before they completed the number adding task is significantly different between the positive and negative feedback conditions in Study 2 but the difference is much smaller and not statistically significant in Study 1: 53.1% vs. 47.4% in the Study 1 (p=0.267, Wilcoxon rank-sum test) and 62.3% vs. 52.8% in Study 2 (p=0.055). The post-task confidence levels which subjects reported after they completed the number adding task and learned their number adding score is significantly different between the feedback conditions in Study 2: subjects reported confidence of 65.5% of being in the top half of their groups when they received positive feedback in the first task compared to the confidence of 44.6% when they received negative feedback (p<0.001). In Study 1 the difference of 3% points in post-task confidence between the positive and negative feedback conditions is

[12] There were two major differences between the two studies: incentivization of beliefs and provision of absolute performance feedback after the first task. The CDFs of elicited confidence beliefs are presented in Figure C1 of the Appendix C. Since the distribution of confidence for the number adding task was different between the non-incentivized and incentivized conditions, we present the results separately for Study 1 and Study 2.

not significant (p=0.666). Finally, the pattern in PR-equivalents is similar in the two studies: subjects revealed £0.13 and £0.18 higher PR-equivalents in Study 1 (p=0.037) and Study 2 (p=0.008) respectively, after receiving positive feedback compared to negative feedback.

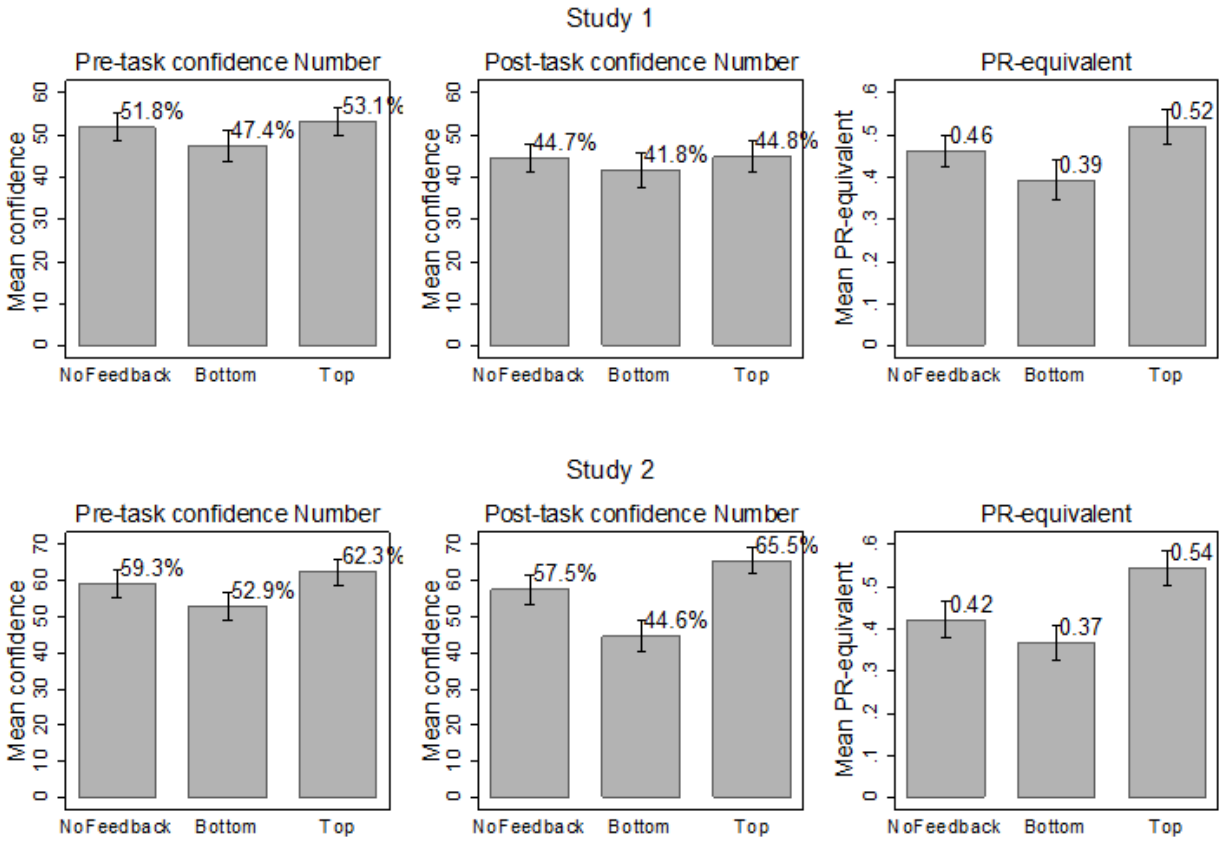


Figure 4: Pre-task and post-task confidence levels and PR-equivalents across feedback conditions. *Error bars show 95% confidence intervals*.

We further examine feedback effects in Tobit regressions under two model specifications in Table 1a and Table 1b. Model 1 controls a number of outcome variables that may mitigate the feedback effects, while Model 2 additionally controls all the variables that we have collected in the experiment, that is, demographics and self-reported measures. We check for the robustness of our results using symmetrically censored least squares[13] and additionally including various specifications for our control variables, such as second and third order polynomials for circle task score and score fixed effects for number adding task score. Results are reported in the Tables C1a and C1b of the Appendix C.

[13] Symmetrically censored least squares is very attractive because it is consistent under very mild assumptions on the distribution of the error terms and is easy to implement. This estimator imposes the symmetric censoring of the dependent variable, which offers some protection against the effects of outlying observations (Santos Silva, 2001).

Table 1a: Feedback effects on confidence and valuations of comparative pay (Study 1)

	Pre-task Confidence (Number)		Post-task Confidence (Number)		PR-equivalent	
Model	(1)	(2)	(1)	(2)	(1)	(2)
Top (vs. no feedback)	2.705	2.480	2.995	2.917	0.101	0.096
	(4.039)	(3.512)	(4.242)	(4.667)	(0.101)	(0.103)
Bottom (vs. no	-6.106	-5.823	-1.102	-1.316	-0.080	-0.062
feedback)	(5.486)	(4.569)	(5.539)	(5.113)	(0.106)	(0.105)
Pre-task Confidence	0.818***	0.786***	0.326**	0.306*	0.001	0.001
(Circle)	(0.107)	(0.093)	(0.156)	(0.158)	(0.001)	(0.001)
Score (Circle)	-0.656	-0.729	-0.233	-0.479	-0.001	-0.003
	(0.556)	(0.492)	(0.637)	(0.592)	(0.008)	(0.008)
Score (Number)	1.864***	1.707***	3.743***	3.750***	0.029***	0.028***
	(0.424)	(0.357)	(0.276)	(0.290)	(0.008)	(0.008)
Top vs. bottom	8.811	8.302	4.097	4.232	0.182	0.158
p-value	[0.054]	[0.015]	[0.578]	[0.574]	[0.026]	[0.088]
Obs.	168	168	168	168	164	164
Feedback only						
Top (vs. bottom)	14.514***	13.949***	7.706	8.174	0.213**	0.181
	(5.080)	(4.554)	(7.830)	(7.978)	(0.107)	(0.124)
Pre-task Confidence	1.028***	0.994***	0.497*	0.503**	0.002	0.003
(Circle)	(0.059)	(0.064)	(0.251)	(0.236)	(0.002)	(0.002)
Score (Circle)	-1.677***	-1.711***	-0.873	-1.089	-0.005	-0.007
	(0.595)	(0.593)	(0.783)	(0.690)	(0.013)	(0.012)
Score (Number)	1.534***	1.464***	3.544***	3.756***	0.032***	0.037***
	(0.553)	(0.495)	(0.324)	(0.359)	(0.010)	(0.009)
Obs.	108	108	108	108	107	107
Demographics & psych. measures	N	Y	N	Y	N	Y

*Standard errors clustered at session level are in parentheses. Demographics include gender and age. Psychological measures include general risk-taking, confidence, and competitiveness. The top vs. bottom estimates and p-values in the first and third panels are from post-estimation Wald tests. * 10%, ** 5%, *** 1% significance levels.*

Table 1a reports results on the sample from both treatments and separately the sample from Feedback treatment of Study 1 while Table 1b reports an identical analysis for Study 2. The key variable of interest is the estimates of Top vs. Bottom, which is either calculated from post-estimation tests when both treatments are analyzed or directly from regressions when only Feedback treatment is analyzed. We find a significant effect of feedback on pre-task confidence in both Study 1 and Study 2: compared to subjects who received negative feedback, those who received positive feedback reported 8.3% to 14.5% points higher level of confidence in Model 1 and 6.7% to 13.9% points higher level of confidence in Model 2. Furthermore, we find some evidence of feedback effects on post-task confidence but only in Study 2. Subjects reported 18.0% points higher level of confidence in Model 1 and 15.6% points higher level of confidence in Model 2 after receiving positive feedback compared to negative feedback. Overall, we conclude that there is some evidence of feedback effects on confidence across tasks.

Table 1b: Feedback effects on confidence and valuations of comparative pay (Study 2)

	Pre-task Confidence (Number)		Post-task Confidence (Number)		PR-equivalent	
Model	(1)	(2)	(1)	(2)	(1)	(2)
Top (vs. no feedback)	1.826	0.450	3.666	2.096	0.113***	0.104***
	(6.030)	(6.312)	(5.497)	(5.385)	(0.035)	(0.030)
Bottom (vs. no	-6.455	-6.227	-8.476	-6.345	-0.066	-0.049
feedback)	(6.580)	(6.291)	(5.521)	(4.596)	(0.073)	(0.065)
Pre-task Confidence	0.557***	0.542***	0.148	0.136*	0.004***	0.004***
(Circle)	(0.100)	(0.094)	(0.094)	(0.076)	(0.001)	(0.001)
Score (Circle)	-0.552***	-0.540**	0.497	0.646	-0.007	-0.006
	(0.205)	(0.223)	(0.535)	(0.462)	(0.005)	(0.007)
Score (Number)	2.152***	2.138***	4.803***	4.975***	0.038***	0.039***
	(0.236)	(0.320)	(0.463)	(0.499)	(0.010)	(0.009)
Top vs. bottom	*8.281*	*6.677*	*12.142*	*8.441*	*0.179*	*0.153*
p-value	*[0.019]*	*[0.057]*	*[0.014]*	*[0.072]*	*[0.004]*	*[0.015]*
Obs.	152	152	152	152	137	137
Feedback only						
Top (vs. bottom)	9.376*	7.680	18.048**	15.575**	0.191***	0.145**
	(5.422)	(5.440)	(7.501)	(6.928)	(0.051)	(0.055)
Pre-task Confidence	0.441***	0.420***	0.114	0.121	0.002*	0.002**
(Circle)	(0.107)	(0.103)	(0.146)	(0.128)	(0.001)	(0.001)
Score (Circle)	-0.829*	-0.737*	-0.490	-0.294	-0.007	-0.003
	(0.483)	(0.433)	(0.943)	(0.907)	(0.009)	(0.012)
Score (Number)	2.545***	2.480***	4.634***	4.635***	0.036**	0.037**
	(0.352)	(0.392)	(1.091)	(1.131)	(0.015)	(0.014)
Obs.	100	100	100	100	86	86
Demographics & psych. measures	N	Y	N	Y	N	Y

*Standard errors clustered at session level are in parentheses. Demographics include gender and age. Psychological measures include general risk-taking, confidence, and competitiveness. The top vs. bottom estimates and p-values in the first and third panels are from post-estimation Wald tests. * 10%, ** 5%, *** 1% significance levels.*

The picture is much clearer for feedback effects on valuations of comparative pay: in both studies, we find a clear difference in revealed PR-equivalents between positive and negative conditions and no feedback condition. In Study 1, we estimate £0.158 to £0.213 higher PR-equivalent after receiving positive feedback than negative feedback. Similarly, in Study 2, we estimate £0.145 to £0.191 higher PR-equivalents. All estimates are statistically significant except in Model 2 of Study 1 where additional controls are added.

4.2 Taste-based or Belief-based Mechanism?

Are the observed feedback effects on PR-equivalents fully accounted for by variations in confidence in the number adding task as predicted by the belief-based mechanism, or does the feedback effect operate by changing subjects' preferences independent of beliefs? To answer this question, in Table 2 we report estimates from regressions of PR-equivalents augmented by both pre-task and post-task confidence of performance in the number adding task.

Here the key variables of interest are Top vs. Bottom and two confidence levels in the number adding task. Compared to the estimates in Table 1, we find that the estimates of Top vs. Bottom are generally smaller by as much as £0.121, but they remain significant in 6 out 8 regressions even after controlling for confidence. This means that the taste-based mechanism plays an important role in our data. The estimates of pre-task and post-task confidence in the number adding task are sometimes significant, especially in Study 2 with Feedback treatment only. Consistent with this observation, the estimates of Top vs. Bottom tend to decrease to a greater extent in Study 2. It suggests that changes in confidence after receiving feedback partially account for the feedback effects on valuations of comparative pay, especially in Study 2 where we incentivized the confidence elicitation. This lends support to the belief-based mechanism. Thus, both the taste-based and belief-based mechanisms appear to explain feedback spillover effects in our data.

To additionally test for mechanism underlying feedback spillover effects we asked the following question in the end-of-study questionnaire of Study 2: "We will be holding different types of experiments in the future. If you had a choice, what kind of experiments would you like to participate in?" The answers ranged from 1 "I would STRONGLY prefer participating in the experiments where my earnings depend on my relative performance compared to other participants' performance" to 5 "I would STRONGLY prefer participating in the experiments where my earnings depend on my own performance NOT compared to other participants' performance" with 3 indicating indifference. We find that the subjects who received negative feedback in the circle task were more averse to participating in experiments with relative performance payment: Mean (SD) of 3.68 (0.198) after receiving negative feedback versus 2.92 (0.178) after receiving positive feedback (p=0.005, Wilcoxon rank-sum test). The difference is significant even after controlling for confidence beliefs in both circle and number adding tasks (Table C2 of the Appendix C). This lends further support to the taste-based mechanism which at least partially operationalizes feedback spillover effects across tasks within the experiment and even spilling over to affect preferences for future experiments.

We should also highlight that in all specifications, the self-reported risk attitude also plays a significant role in explaining PR-equivalents. All estimates of risk attitude are at least significant at the 10% level. This is not surprising given that valuations of comparative pay, by design, is related to both confidence and risk attitude. Our data, however, suggests that relative performance feedback from an unrelated task can operate on its own, plausibly through a taste-based mechanism, to affect individuals' valuations of comparative pay in addition to variations in confidence and risk attitude.

Table 2: Feedback effects on PR-equivalents are independent of feedback effects on confidence

	PR-equivalent			
	Study 1		Study 2	
Top (vs. no feedback)	0.090	0.088	0.103***	0.099***
	(0.093)	(0.098)	(0.024)	(0.020)
Bottom (vs. no feedback)	-0.065	-0.051	-0.019	-0.013
	(0.097)	(0.099)	(0.054)	(0.056)
Pre-task confidence (Circle)	-0.002	-0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Score (Circle)	0.001	-0.002	-0.006	-0.005
	(0.008)	(0.008)	(0.005)	(0.006)
Pre-task confidence (Number)	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Post-task confidence	0.003	0.002	0.003**	0.002
(Number)	(0.002)	(0.009)	(0.001)	(0.001)
Score (Number)	0.016	0.018*	0.022*	0.026**
	(0.010)	(0.009)	(0.012)	(0.011)
Top vs. bottom	*0.155*	*0.138*	*0.123*	*0.113*
p-value	*[0.022]*	*[0.083]*	*[0.018]*	*[0.049]*
Obs.	164	164	137	137
Feedback only				
Top vs. bottom	*0.193***	*0.166*	*0.054**	*0.024*
	(0.094)	*(0.106)*	*(0.032)*	*(0.035)*
Pre-task confidence (Circle)	0.001	0.002	-0.000	-0.000
	(0.004)	(0.004)	(0.001)	(0.001)
Score (Circle)	-0.002	-0.005	0.001	0.004
	(0.014)	(0.013)	(0.010)	(0.011)
Pre-task confidence (Number)	0.000	0.000	0.004***	0.005***
	(0.003)	(0.003)	(0.001)	(0.001)
Post-task confidence	0.003	0.002	0.004***	0.003***
(Number)	(0.002)	(0.002)	(0.001)	(0.001)
Score (Number)	0.023*	0.030**	0.010	0.015
	(0.013)	(0.012)	(0.014)	(0.012)
Obs.	107	107	86	86
Demographics & psych. measures	N	Y	N	Y

*Standard errors clustered at session level are in parentheses. Demographics include gender and age. Psychological measures include general risk-taking, confidence, and competitiveness. The top vs. bottom estimates and p-values in the first panel are from post-estimation Wald tests. * 10%, ** 5%, *** 1% significance levels.*

4.3 Gendered feedback spillover effects

Past research shows that men and women have different behavioral responses to previous competition outcomes (Buser, 2016; Gill & Prowse, 2014). Since we used the number adding task, the often-used task to study gender differences in willingness to compete, as an additional analysis we also explore whether there are gender differences in valuations of comparative pay and whether feedback spillover effects interact with gender.

In Table 3, we analyze the gender differences in PR-equivalents controlling for other observable variables and find that gender differences in the absolute level of PR-equivalents are

influenced by feedback. The regression results show that the average PR-equivalent of women is £0.116 to £0.145 lower than that of men in Study 1 and £0.134 to £0.157 in Study 2, although the latter estimates are not precise. Importantly, the interaction of the female dummy with the Feedback treatment suggests that receiving feedback significantly increases the average PR-equivalents of women compared to men by £0.288 to £0.321 in Study 1 and by £0.188 to £0.239 in Study 2. Receiving feedback decreases (or reverses) the gender gap in valuations of comparative pay in both studies, independently of variations in confidence and risk attitude.

Table 3: Feedback decreases (or reverses) the gender gap in valuations of comparative pay

	PR-equivalent			
	Study 1		Study 2	
Female	-0.145***	-0.116***	-0.134	-0.157
	(0.035)	(0.040)	(0.094)	(0.100)
Female×Feedback	0.288***	0.321***	0.188	0.239*
	(0.098)	(0.092)	(0.124)	(0.132)
Feedback	-0.144	-0.149	-0.046	-0.065
	(0.109)	(0.099)	(0.059)	(0.065)
Pre-task confidence (Number)	0.002**	0.001	0.004**	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)
Post-task confidence (Number)	0.003	0.002	0.003**	0.002
	(0.002)	(0.002)	(0.001)	(0.001)
Score (Number)	0.016*	0.018*	0.022**	0.027**
	(0.008)	(0.008)	(0.011)	(0.011)
Obs.	164	164	137	137
Demographics & psych. measures	N	Y	N	Y

*Standard errors clustered at session level are in parenthesis. Demographics include age. Psychological measures include general risk-taking, confidence, and competitiveness. Excludes subjects who switched non-monotonically. * 10%, ** 5%, *** 1% significance levels.*

What are the forces driving this result? We look at whether there are gender differences in feedback effects on valuations of comparative pay. Table 4 reports the estimates separately for different genders. We find, compared to no feedback, that men who receive negative feedback reduce their valuations of comparative pay, whereas women who receive positive feedback increase their valuations. The effects are qualitatively similar in both studies, but are only statistically significant in Study 1. In fact, if we only focus on the sign of the estimates, men almost always reduce their valuations of comparative pay regardless of which feedback they receive (except in Study 2 when they receive positive feedback). In contrast, women increase their valuations regardless of which feedback they receive. This result is virtually unchanged even after controlling for the two confidence measures in the number adding task. This suggests that the gender difference in feedback effects are mainly driven by changes in tastes for

comparisons rather than by changes in confidence. One conjecture is that men have relatively higher tastes for comparisons than women to start with. Relative performance feedback might partially satisfy men's need for comparisons, while stimulating women's need for comparisons. Testing for the robustness of this finding and offering fuller explanations are beyond the scope of this paper and will be left for future research.

Table 4: Gender differences in feedback effects on valuations of comparative pay

	PR-equivalent			
	Study 1		Study 2	
	Men	Women	Men	Women
Top (vs. no feedback)	-0.105	0.250**	0.019	0.153
	(0.119)	(0.108)	(0.073)	(0.106)
Bottom (vs. no feedback)	-0.220**	0.117	-0.139*	0.062
	(0.083)	(0.140)	(0.079)	(0.082)
Pre-task confidence (Circle)	0.004***	-0.001	0.001	0.005***
	(0.001)	(0.002)	(0.001)	(0.002)
Score (Circle)	-0.008	-0.002	-0.000	-0.008
	(0.012)	(0.008)	(0.010)	(0.006)
Score (Number)	0.017	0.018*	0.037***	0.045***
	(0.010)	(0.009)	(0.010)	(0.010)
Obs.	74	90	74	63
Top (vs. no feedback)	-0.105	0.242**	-0.011	0.159
	(0.113)	(0.095)	(0.068)	(0.108)
Bottom (vs. no feedback)	-0.180**	0.087	-0.126	0.136
	(0.077)	(0.134)	(0.083)	(0.082)
Pre-task confidence (Circle)	0.001	-0.003	-0.002	0.004*
	(0.003)	(0.004)	(0.002)	(0.002)
Score (Circle)	-0.005	0.002	0.003	-0.008
	(0.011)	(0.014)	(0.009)	(0.007)
Pre-task confidence (Number)	0.003	0.001	0.004	0.003**
	(0.002)	(0.002)	(0.003)	(0.002)
Post-task confidence (Number)	0.001	0.004*	0.002	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Score (Number)	0.009	0.030*	0.028*	0.018***
	(0.014)	(0.016)	(0.014)	(0.005)
Obs.	74	90	74	63
Demographics & psych. measures	Y	Y	Y	Y

*Standard errors clustered at session level are in parentheses. Demographics include age. Psychological measures include general risk-taking, confidence, and competitiveness. Excludes subjects who switched non-monotonically. * 10%, ** 5%, *** 1% significance levels.*

5 Conclusion

Past research on feedback effects has shown the importance of the informational value of feedback, which corrects biased beliefs and mitigates competitive entry failures within single

task domains. In this paper, we investigate the effects of relative performance feedback in a setup where workers have to work on unrelated tasks. Using a laboratory experiment we examine whether relative performance feedback from the first task affects confidence and preference for (i.e. valuations of) comparative pay in a different second task. We thus depart from most of the existing literature by focusing on the interim effects of feedback across two different tasks rather than within the same task.

We present a novel finding that feedback from the first task significantly affects preferences for comparative pay in the second unrelated task. We also find that this effect can operate independently of variations in confidence. The results underline the importance of the psychological value of feedback in addition to its widely studied informational value: according to our results, relative performance feedback that employees receive in one job task can affect their preferences independent of its effects on their confidence and beliefs about subsequent unrelated job tasks. More importantly, this suggests that a taste-based mechanism, besides the often-discussed belief-based mechanism, can explain the feedback effects on preferences of comparative pay.

Another novel finding is the differential impacts of feedback on women's and men's preferences for comparative pay. Previous literature has stressed that feedback helps mitigate or eliminate gender differences in willingness to compete via correcting biased beliefs of relative performance. For example, Wozniak et al. (2014, 2015) show that giving out relative performance feedback in number adding tasks encourages high-skilled women to and discourages low-skilled men from entering competitions. These studies show that feedback helps close the gender gap in competitive entry behavior, and consequently improves economic efficiency through sorting by ability. Thus, feedback is argued to be an alternative and perhaps a more efficient affirmative action policy compared to other policies studied in the literature (Balafoutas & Sutter, 2012; Niederle, et al., 2013). In our setup, we diminish and control the informational value of feedback by assessing feedback effects across two different tasks and yet show that feedback affects gender differences in valuations of comparative pay. Our result suggests the possibility of a taste-based rather than a belief-based mechanism by which feedback operates to align or even reverse men and women's preferences for comparative pay. A deeper understanding of this result is left for future research.

In practical applications of our results, organizations should be aware of the context and timing in which tasks are assigned to workers. For example, by providing easy tasks at the beginning may improve success in subsequent tasks that require high confidence and competitiveness, such as leading negotiations, pitching new ideas to potential investors, and

performing in high accuracy sports tournaments such as golf (Rosenqvist & Skans 2015). Similarly, schools can create positive environments with more positive feedback to foster confidence and competitiveness of students which are predictors of students' career choices (Buser, Niederle, & Oosterbeek, 2014; Schulz & Thöni, 2016). This may prove especially important for encouraging students from disadvantaged backgrounds to pursue competitive careers, such as those from ethnic and racial minority groups and women. Another interesting outlet to use feedback spillovers across different tasks are jobcentres: anecdotal evidence suggests that the psychometric and ability tests that jobseekers fill at UK jobcentres are calibrated so that the test results are always positive and give some positive feedback to jobseekers to improve their confidence (Malik, 2013) and thus the probability of successful job search. Our results provide (partial) support for such policies and emphasize the importance of psychological value of feedback.

One concern about the generalizability of our results is that all participants in our experiment are university students. If we take into account that students will become employees of public and private sectors, their behavior and preferences are of particular interest to researchers. Another concern is that lab experiments examine preferences and behavior in artificial environments with limited external validity. However, we anticipate that the observed feedback effects across two clearly dissimilar laboratory tasks could have underestimated the true extent of feedback effects in the field where the boundaries between distinct job tasks are more blurred. Moreover, a recent study by Buser, Niederle, & Oosterbeek (2014) shows that competitive preferences measured in the laboratory strongly predict students' academic track and career choices. Extrapolating their results to our experiment, our findings suggest that the long-term consequences of feedback on students' willingness to pursue more competitive academic tracks and consequently high-paid careers in the future may have psychological roots in earlier path-contingent yet unrelated events.

As far as we know, our study presents the first piece of evidence on feedback spillover effects between different task domains and disentangling between informational and taste-based effects of feedback. The natural step for future research is to investigate whether significant spillover effects across tasks are present in other economically relevant variables such as performance under flat rate and competitive incentive schemes,[14] and willingness to compete in

[14] Perhaps not surprisingly we don't find feedback effects on piece rate performance in our experiment. This resonates with the mixed findings in literature: while some researchers find that ranking feedback without extra financial rewards improves performance (Azmat & Iriberri, 2010, 2016; Blanes i Vidal & Nossol, 2011; Hannan, Krishnan, & Newman, 2008), others find the opposite effects (Barankay, 2012a, 2012b) or no effects (Eriksson, Poulsen, & Villeval, 2009). When working under flat-rate incentives, individuals who directly observe others'

future tasks. Of particular interest for policy makers is whether any of these effects persist in real world organizations and other naturally occurring settings.

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, *101*(2), 470–492.

Ahn, T. K., Ostrom, E., Schmidt, D., Shupp, R., & Walker, J. (2001). Cooperation in PD Games: Fear, Greed, and History of Play. *Public Choice*, *106*(1/2), 137–155.

Arya, A., & Mittendorf, B. (2004). Using Job Rotation to Extract Employee Information. *Journal of Law Economics, & Organization*, *20*(2), 400–414.

Azmat, G., & Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, *94*, 435–452.

Azmat, G., & Iriberri, N. (2016). The Provision of Relative Performance Feedback: An Analysis of Performance and Satisfaction. *Journal of Economics & Management Strategy*, *25*(1), 77–110.

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

Barankay, I. (2012a). *Rank Incentives: Evidence from a randomized workplace experiment*. Working paper.

Barankay, I. (2012b). *Rankings and Social Tournaments: Evidence from a Crowd-Sourcing Experiment*. Working paper.

Benoit, J.-P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica*, *79*(5), 1591–1625.

Benoit, J.-P., Dubra, J., & Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association*, *13*(2), 293–329.

Blanes i Vidal, J., & Nossol, M. (2011). Tournaments Without Prizes: Evidence from Personnel Records. *Management Science*, *57*(10), 1721–1736.

Boring, A., & Brown, J. (2016). Gender, Competition and Choices in Higher Education. *Working Paper*. Retrieved from http://www.anneboring.com/higher-education.html

performance (Falk & Ichino, 2006; Mas & Moretti, 2009) or expect to receive relative performance feedback (Cadsby, et al., 2015; Gerhards & Siemer, 2015; Hannan, et al., 2013; Kosfeld & Neckermann, 2011; Kuhnen & Tymula, 2012) generally improve their performance, though specific ranks have differential impacts (Gill et al., 2016). When tournament incentives are applied, similar mixed evidence emerges: Eriksson et al. (2009) find motivational effects of feedback, whereas Hannan et al. (2008) find de-motivational effects. Most of these studies focus on the interim effects of feedback on performance within a task rather than feedback effects on performance in a subsequent task.

Brown, J., & Minor, D. B. (2014). Selecting the Best? Spillover and Shadows in Elimination Tournaments. *Management Science*, *60*(12), 3087–3102.

Bruhin, A., Goette, L., Haenni, S., & Jiang, L. (2015). Spillovers of Prosocial Motivation: Evidence from an Intervention Study on Blood Donors. CEPR Discussion Paper No. DP10345.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2013). Overconfidence and Social Signalling. *Review of Economic Studies*, *80*(3), 949–983.

Buser, T. (2016). The Impact of Losing in a Competition on the Willingness to Seek Further Challenges. *Management Science*, *62*(12), 3439–3449.

Buser, T., Gerhards, L., & van der Weele, J. J. (2017). Measuring Responsiveness to Feedback as a Personal Trait. *Forthcoming in the Journal of Risk and Uncertainty*.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices. *Quarterly Journal of Economics*, *129*(3), 1409–1447.

Cadsby, C. B., Engle-Warnick, J., Fang, T., & Song, F. (2015). *Psychological Incentives, Financial Incentives, and Risk Attitudes in Tournaments: An Artefactual Field Experiment*. IZA Discussion paper No. 9565.

Cason, T. N., Savikhin, A. C., & Sheremeta, R. M. (2012). Behavioral spillovers in coordination games. *European Economic Review*, *56*, 233–245.

Coviello, B. D., Ichino, A., & Persico, N. (2014). Time Allocation and Task Juggling. *American Economic Review*, *104*(2), 609–623.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, *9*(3), 522–550.

Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, *47*, 1–16.

Dreber, A., von Essen, E., & Ranehill, E. (2014). Gender and competition in adolescence: task matters. *Experimental Economics*, *17*(1), 154–172.

Eberlein, M., Ludwig, S., & Nafziger, J. (2011). The Effects of Feedback on Self-assessment. *Bulletin of Economic Research*, *63*(2), 177–199.

Ederer, F. (2010). Feedback and Motivation in Dynamic Tournaments. *Journal of Economics and Management Strategy*, *19*(3), 733–769.

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*, 114–138.

Eriksson, T., Poulsen, A., & Villeval, M. C. (2009). Feedback and Incentives: Experiment Evidence. *Labour Economics*, *16*, 679–688.

Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior and Organization*, *80*, 532–545.

Ertac, S., & Szentes, B. (2011). *The Effect of Information on Gender Differences in Competitiveness: Experimental Evidence*. Koc University Economic Research Forum Working Paper 1104.

Falk, A., & Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, *24*(1), 39–57.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Fishbach, A., Eyal, T., & Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Social and Psychology Compass*, *4*(8), 517–530.

Gächter, S., Huang, L., & Sefton, M. (2016). Combining "real effort" with induced effort costs: the ball-catching task. *Experimental Economics*, *19*(4), 687–712. https://doi.org/10.1007/s10683-015-9465-9

Gautier, A., & Wauthy, X. (2007). Teaching versus research: A multi-tasking approach to multi-department universities. *European Economic Review*, *51*(2), 273–295.

Gerhards, L., & Siemer, N. (2016). THE IMPACT OF PRIVATE AND PUBLIC FEEDBACK ON WORKER PERFORMANCE-EVIDENCE FROM THE LAB. *Economic Inquiry*, *54*(2), 1188–1201.

Gill, D., Kissova, Z., Lee, J., & Prowse, V. (2016). *First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision*. IZA Discussion paper No. 9286.

Gill, D., & Prowse, V. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *American Economic Review*, *102*(1), 469–503.

Gill, D., & Prowse, V. (2014). Gender differences and dynamics in competition: The role of luck. *Quantitative Economics*, *5*, 351–376.

Hannan, L., Krishnan, R., & Newman, A. H. (2008). The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *Accounting Review*, *83*(4), 893–913.

Hannan, L., McPhee, G. P., Newman, A. H., & Tafkov, I. D. (2013). The Effect of Relative Performance Information on Performance and Effort Allocation in a Multi-Task

Environment. *Accounting Review*, *88*(2), 553–575.

Hollard, G., Massoni, S., & Vergnaud, J.-C. (2016). In search of good probability assessors: an experimental comparison of elicitation rules for confidence judgments. *Theory and Decision*, *80*(3), 363–387.

Hsieh, A.-T., & Chao, H.-Y. (2004). A reassessment of the relationship between job specialization, job rotation and job burnout: example of Taiwan's high-technology industry. *International Journal of Human Resource Management*, *15*(6), 1108–1123.

Ifcher, J., & Zarghamee, H. (2016). Pricing competition: a new laboratory measure of gender differences in the willingness to compete. *Experimental Economics*, *19*(3), 642–662.

Johlke, M. C. (2006). Sales presentation skills and salesperson job performance. *Journal of Business & Industrial Marketing*, *21*(5), 311–319.

Kamas, L., & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, *83*(1), 82–97.

Klumpp, T., & Polborn, M. K. (2006). Primaries and the New Hampshire Effect. *Journal of Public Economics*, *90*, 1073–1114.

Knez, M., & Camerer, C. (2000). Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games. *Organizational Behavior and Human Decision Processes*, *82*(2), 194–216.

Kosfeld, M., & Neckermann, S. (2011). Getting More Work for Nothing? Symbolic Awards and Worker Performance. *American Economic Journal: Microeconomics*, *3*(3), 86–99.

Kőszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, *4*(4), 673–707.

Kuhnen, C. M., & Tymula, A. (2012). Feedback, Self-Esteem, and Performance in Organizations. *Management Science*, *58*(1), 94–113.

Malik, S. (2013). Jobseekers made to carry out bogus psychometric tests | Society | The Guardian. https://doi.org/www.theguardian.com/society/2013/apr/30/jobseekers-bogus-psychometric-tests-unemployed

Malueg, D. A., & Yates, A. J. (2010). Testing Contest Theory: Evidence from Best-of-Three Tennis Matches. *Review of Economics and Statistics*, *92*(3), 689–692.

Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, *99*(1), 112–145.

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). *Managing self-confidence: Theory and experimental evidence*. NBER Working Paper No. 17014.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and*

Human Decision Processes, *103*(2), 197–213.

Murad, Z. (2016). Does group discussion lead to better informed and more strategic market entry decisions? *Economics Letters*, *144*, 25–28.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, *122*(3), 1067–1101.

Pohlhaus, J. R., Jiang, H., Wagner, R. M., Schaffer, W. T., & Pinn, V. W. (2011). Sex Differences in Application, Success, and Funding Rates for NIH Extramural Programs. *Academic Medicine*, *86*(6), 759–767.

Rose, J. P., & Windschitl, P. D. (2008). How egocentrism and optimism change in response to feedback in repeated competitions. *Organizational Behavior and Human Decision Processes*, *105*(2), 201–220.

Rosenqvist, O., & Skans, O. N. (2015). Confidence enhanced performance? – The causal effects of success on future performance in professional golf tournaments. *Journal of Economic Behavior & Organization*, *117*, 281–295.

Schottner, A. (2008). Relational contracts, multitasking, and job design. *Journal of Law, Economics, and Organization*, *24*(1), 138–162.

Schulz, J. F., & Thöni, C. (2016). Overconfidence and Career Choice. *PLOS ONE*, *11*(1), e0145126.

Santos Silva, J. M. C. S. (2001). Influence Diagnostics and Estimation Algorithms for Powell's SCLS. *Journal of Business & Economic Statistics*, *19*(1), 55–62.

van Veldhuizen, R. (2016). *Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness?* WZB Discussion Paper. Retrieved from https://www.econstor.eu/handle/10419/148346

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choice. *Journal of Labor Economics*, *32*(1), 161–198.

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2015). *Gender differences: The effects of free and costly feedback on the willingness to compete*. Available at SSRN: http://ssrn.com/abstract=1976073.

Zeelenberg, M., Beattie, J., van der Pligt, J., & de Vries, N. K. (1996). Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making. *Organizational Behavior and Human Decision Processes*, *65*(2), 148–158.

Appendices (intended for online publication)

Appendix A

A1 Experimental Instructions for the Main Experiment

General Information

Welcome! You are about to participate in an experiment on decision making. You will receive £2 for showing up to and £5 for completing this experiment. Additionally, you can earn money during the experiment.

The experiment will consist of several parts. At the end of the experiment, one of the parts will be randomly selected by the computer. Your final earnings will be determined according to your performance in the selected part. The instructions on how your performance is determined for each part will be shown on your screens. Please read the instructions carefully and take notes if necessary as you will not be able to go back to read them again when the experiment starts.

You may leave at any point during the experiment, if you do not wish to complete the experiment. If you leave the experiment before it is completed, you will only be paid the show up fee of £2.

Throughout the experiment you must not communicate with other participants. The use of any electronic devices is strictly prohibited. If you break these rules, you will be excluded from the experiment without receiving any fee. Please make sure that all such devices are turned off and put away out of sight.

Your decisions are anonymous and under no circumstances will be linked to your identity.

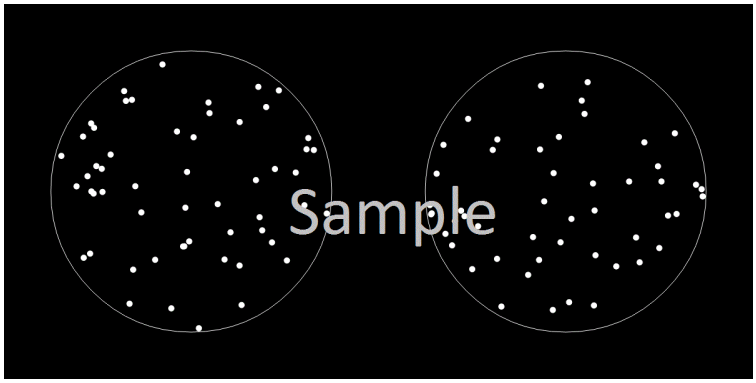
If you have any questions at any point during the experiment, please raise your hand and someone will come to your desk to answer it. If you agree by these rules, you can sign and date the CONSENT FORM on your desks that you are willing to participate in this experiment and consenting the use of your data in the final report of the project.

[On screen instructions]

Part 1

In this part, you will complete the Circle Task. In the Circle Task, you will be asked to pick a circle with more dots in a series of circle pairs. One of the circles will contain 50 dots and the other one 55 dots. The figure below shows an example of a circle pair. Each circle pair will appear on your screens for 1 second only. After the circle pair disappears from your screen, you will be asked to judge whether the right or the left circle contained more dots. You have to indicate your judgement by clicking on the "Left" or "Right" button. When you click the button of your choice, you will move to the next pair of circles. The task consists of 60 circle pairs. Your performance score in this task will be the total number of correctly judged circles. **If Part 1 is selected for the payment, you will earn £0.20 per each correct answer.**

During the experiment you will be matched with three other participants to form a group consisting of 4 participants. You will be in the same group throughout the whole course of the experiment. All of your group members will also complete the Circle task and receive a performance score based on the number of correctly judged circles.



Press to Continue

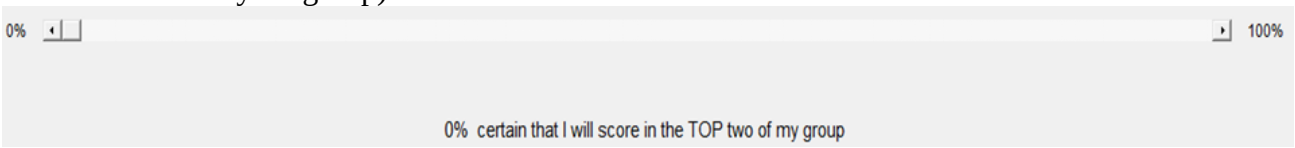
[Study 1]

Before starting the Circle task, please indicate the probability that your performance score will be one of the top two scores of your group. Your answer will not affect your earnings in any way, but please try to express your beliefs of scoring in the top two of your group as accurately as you can.

[Feedback treatment] At the end of the task, you will be informed what your performance score was and whether your score was in the top two or bottom two of your group.

[NoFeedback treatment] At the end of the task, you will be informed what your performance score was. At the end of the experiment, you will also be informed whether your score was in the top two or bottom two of your group.

On a scale from 0 to 100, please indicate the probability that your score will be in the TOP two of your group. (Please assign 0% if you are completely certain you will score in the BOTTOM two and 100% if you are completely certain you will score in the TOP two of your group. Assign intermediary values if you are uncertain whether you will score in the TOP or BOTTOM two of your group).



Press to Continue

Feedback after Part 1:

[Feedback treatment] This is the end of Part 1. Your performance score for the Circle task was [#] correctly judged circles out of 60. Your score was one of the [Top or Bottom] two scores of your group.

[NoFeedback treatment] This is the end of Part 1. Your performance score for the Circle task was [#] correctly judged circles out of 60.

Press to Continue

[Study 2]

Before starting the Circle task, please indicate the probability that your performance score will be one of the top two scores of your group. The accuracy of your stated probability that your performance will be in the top two scores of your group will influence your payment. The experiment is designed such that you have the highest chance to earn money if you state the true probability with which you believe that your score is in the top two scores in your group.

Your stated probability will determine your choice between two options:

Option A: you receive an additional £1 if your performance score in the Circle task is in the top two scores in your group of 4 participants

Option B: you receive an additional £1 with probability X

X is a number between 0% and 100% and will be randomly picked by the computer. Each number between 0% and 100% is equally likely to be picked. Your reported probability of being in the top two determines the choice between option A and option B in the following way. If X is higher than your reported probability, option B is chosen and you receive £1 with probability X. If X is lower than your reported probability, option A is chosen and you receive £1 if your performance score is in the top two scores of your group.

This means that you always maximize the chances of winning the £1 if you state the true probability with which you believe that your score will be in top two. We can illustrate this with an example.

For example, let's assume that after reading the instructions for the Circle task you believe that it is 62% likely that your score will be in top two scores of your group.

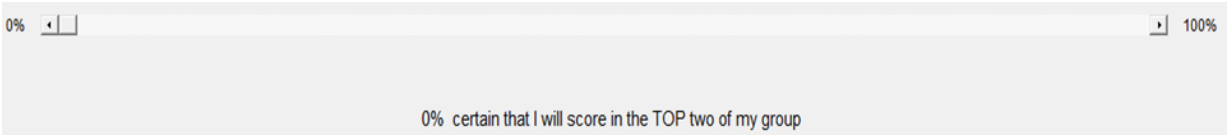
- Suppose you report your belief accurately. Then, if X is below 62% you will get option A, which gives you a higher probability of winning (namely 62%) than you would get with option B (namely X). If X is above 62% you will get option B, which gives you a higher chance of winning (namely X) than you would get with option A (namely 62%). So, whatever the value of X, you will always get the option that gives you the highest chance of winning.
- Suppose you report an inaccurate belief, let's say 10%. Then, if X is below 10% you will get option A, which gives you a higher probability of winning (namely 62%) than you would get with option B (namely X). If X is above 10%, you will get option B, but this option may give you a lower chance of winning than option A. If X is between 10% and 62%, you get option B, even though you are better off with option A which gives you a probability of 62%. Reporting a wrong belief will therefore reduce your chances of winning the £1.

The logic in this example holds for all probabilities beliefs. **The mechanism may look very complicated but whatever your belief is that your score will be in the top two scores of your group of 4 participants, you maximize your chance of winning £1 if you report your true belief. We will add £1 to your total earnings in Part 1 if you win.**

[Feedback treatment] At the end of the task, you will be informed whether your score was in the top two or bottom two of your group.

[NoFeedback treatment] At the end of the experiment, you will be informed whether your score was in the top two or bottom two of your group.

Before you start the Circle task, on a scale from 0 to 100, please indicate the probability that your score will be in the TOP two of your group. You are most likely to win £1 if you are as accurate as possible. (Please assign 0% if you are completely certain you will score in the BOTTOM two and 100% if you are completely certain you will score in the TOP two of your group. Assign intermediary values if you are uncertain whether you will score in the TOP or BOTTOM two of your group).



Press to Continue

Feedback After Part 1:

[Feedback treatment] This is the end of Part 1. Your performance score for the Circle task was one of the [Top or Bottom] two scores of your group.

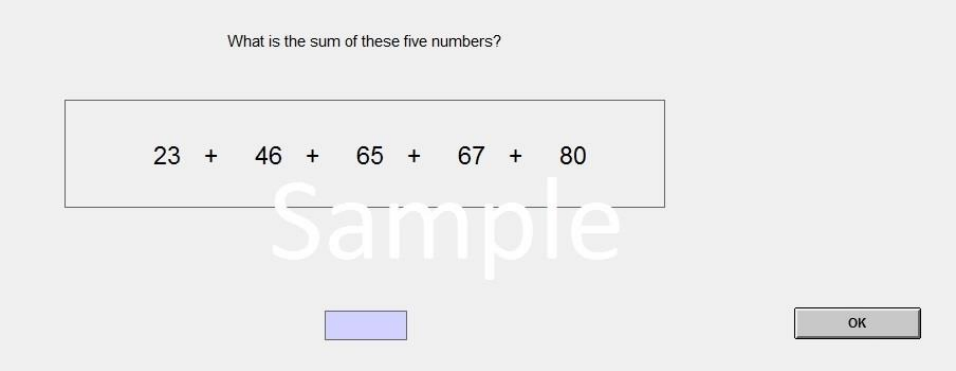
[NoFeedback treatment] This is the end of Part 1.

Press to Continue

Part 2

In this part, you will complete the Number Adding Task. In the Number Adding task, you will be asked to add five two digit numbers in a series of tables. The figure below shows the work screen you will use for this task. You will enter the answer into the box below the table. After you have entered the answer, you can click the NEXT button. No matter whether the answer is correct or not, a new table will be generated. Your performance score in this task will be the number of correct answers at the end of the 300 seconds. **If Part 2 is selected for the payment, you will earn £0.50 per each correctly solved number addition.** If you enter a wrong answer, you will earn nothing for that table. You cannot use a calculator to determine the answer, however you are allowed to use the pen and paper on your desks to help you with your calculations.

You are in the same group with the same participants as you were in Part 1. All of your group members will complete the Number Adding task and receive a performance score based on the number of correctly solved number additions.



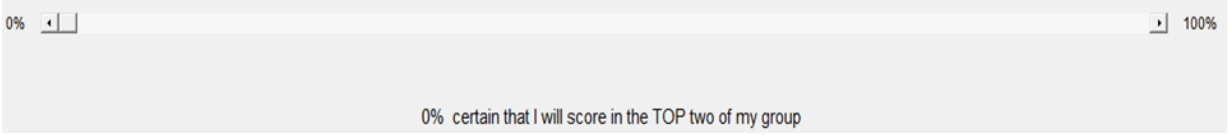
Press to Continue

[Study 1]

Before starting the Number Adding task, please indicate the probability that your performance score will be one of the top two scores of your group. Your answer will not affect your earnings in any way, but please try to express your beliefs of scoring in the top two of your group as accurately as you can.

At the end of the task, you will be informed what your performance score was. At the end of the experimental session, you will also be informed whether your score was in the top two or bottom two of your group.

On a scale from 0 to 100, please indicate the probability that your score will be in the TOP two of your group. (Please assign 0% if you are completely certain you will score in the BOTTOM two and 100% if you are completely certain you will score in the TOP two of your group. Assign intermediary values if you are uncertain whether you will score in the TOP or BOTTOM two of your group).



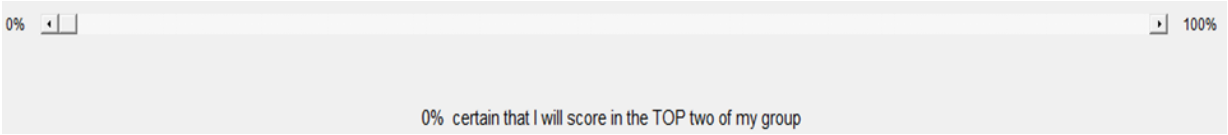
Press to Continue

[Study 2]

Before starting the Number Adding task, please indicate the probability that your performance score will be one of the top two scores of your group. The accuracy of your stated probability that your performance will be in the top two scores of your group will influence your payment in the same way as before. You have the highest chance to win an additional £1 to your Part 2 earnings if you state the true probability with which you believe that your score is in the top two scores in your group.

At the end of the Number Adding task, you will be informed what your performance score was. At the end of the experimental session, you will also be informed whether your score was in the top two or bottom two of your group.

Before you start the Number Adding task, on a scale from 0 to 100, please indicate the probability that your score will be in the TOP two of your group. You are most likely to win £1 if you are as accurate as possible. (Please assign 0% if you are completely certain you will score in the BOTTOM two and 100% if you are completely certain you will score in the TOP two of your group. Assign intermediary values if you are uncertain whether you will score in the TOP or BOTTOM two of your group).



Press to Continue

[Study 1]

Feedback after Part 2:

This is the end of Part 2. Your performance score for the Number Adding task was [#] correctly solved number additions. Now that you know your score, please indicate the probability that your score in Part 2 was one of the top two scores of your group. Your answer will not affect your earnings in any way, but please try to express your beliefs of scoring in the top two of your group as accurately as you can. At the end of experiment, you will be informed whether your score was in the TOP or BOTTOM two scorers of your group.

On a scale from 0 to 100, please indicate the probability that your score was in the TOP two of your group? (Please assign 0% if you are completely certain you scored in the BOTTOM two of your group and 100% if you are completely certain you scored in the TOP two of your group. Assign intermediary values if you are uncertain whether you scored in the TOP or BOTTOM two of your group).

Press to Continue

[Study 2]

Feedback After Part 2:

This is the end of Part 2. Your performance score for the Number Adding task was [#] correctly solved number additions. Now that you know your score, please indicate the probability that your score in Part 2 was one of the top two scores of your group. The accuracy of your stated probability that your performance will be in the top two scores of your group will influence your payment in the same way as before. You have the highest chance to win an additional £1 to your Part 3 earnings if you state the true probability with which you believe that your score is in the top two scores in your group.

At the end of experiment, you will be informed whether your score was in the TOP or BOTTOM two scorers of your group.

On a scale from 0 to 100, please indicate the probability that your score was in the TOP two of your group. You are most likely to win £1 if you are as accurate as possible. (Please assign 0% if you are completely certain you scored in the BOTTOM two of your group and 100% if you are completely certain you scored in the TOP two of your group. Assign intermediary values if you are uncertain whether you scored in the TOP or BOTTOM two of your group).

Press to Continue

Part 3

In this part, you do not have to complete any task. Instead you will choose how you want to be paid based on the performance score you received in Part 2, which was the number of correct solved number additions. In the table below, you will choose to be paid either according to Piece Rate or Tournament. You will have to make a choice for each row of the table between Piece Rate (ranging from £0.00 per correctly solved number addition to £1.00 per correct solved number addition) and Tournament.

If Part 3 is selected for the payment, one of the ten rows will be randomly chosen and you will be paid according to the choice you made on that row.

If you choose Tournament in the selected row, your performance will be evaluated relative to other members of your group. If you scored in the top two of your group, you will earn £1.00 per each correctly solved number addition. If you scored in the bottom two of your group, you will earn £0.00. Ties in the scores will be randomly resolved by the computer.

If you choose Piece Rate on the selected row, you will earn according to the piece rate of that row for each correctly solved number addition.

Below you can make your choice whether you want to be paid by Piece Rate or Tournament for each row of the table. If you have any questions, please raise your hand and someone will come to your desk to answer it.

Row	A: Tournament	Your Choice	B: Piece Rate
10	You get £1.00 per correctly solved number addition if your score in Part 2 was in the TOP two scores of your group and £0.00 if your score in Part 2 was in the Bottom two scores of your group.	□ □	£1.00 per correct number addition
9		□ □	£0.90 per correct number addition
8		□ □	£0.80 per correct number addition
7		□ □	£0.70 per correct number addition
6		□ □	£0.60 per correct number addition
5		□ □	£0.50 per correct number addition
4		□ □	£0.40 per correct number addition
3		□ □	£0.30 per correct number addition
2		□ □	£0.20 per correct number addition
1		□ □	£0.10 per correct number addition

Questionnaire

End of Part 3: This is the end of the experiment. Before continuing to the final feedback and payment stages, please answer the following questions as accurately as you can. Your answers are anonymous and will not be linked to your identity.

[Study 2]: We will be holding different types experiments in the future. If you had a choice, what kind of experiments would you like to participate in?

1. I would STRONGLY prefer participating in the experiments where my earnings depend on my relative performance compared to other participants' performance.
2. I would SLIGHTLY prefer participating in experiments where my earnings depend on my relative performance compared to other participants' performance.
3. I am completely indifferent between the types of experiment.
4. I would SLIGHTLY prefer participating in the experiments where my earnings depend on my own performance NOT compared to other participants' performance.
5. I would STRONGLY prefer participating in the experiments where my earnings depend on my own performance NOT compared to other participants' performance.]

1. What is your gender? Male □ Female □
2. What is your age? _____
3. What is your nationality?
   British □ Other □
4. If you are a student, what is your subject area? _____
5. On a scale of 1 to 7, how willing are you to take risks in general?
   1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □
   Not at all willing Very willing

6. On a scale of 1 to 7, how confident are you as a person?
   1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □
   Not at all confident Extremely confident

7. On a scale of 1 to 7, how competitive are you as a person?
   1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □
   Not at all competitive Extremely Competitive

8. Do you think that the performance in the Circle task and Number Adding task are related? (that is, people who tend to score high in the Circle Task also score high in the Number Adding task, and vice versa.)
   Yes □ No □ Don't Know □

End of Study Feedback Screen

In Part 1, your performance score in the Circle task was [#] correctly judged circles and you were in the [Top or Bottom] two of your group.
In Part 2, your performance score in the Number Adding task was [#] correctly solved number additions and you were in the [Top or Bottom] two of your group.
Part [1, 2 or 3] was chosen for your payment.
You will be paid a show-up fee of £2 and a participation fee of £5 in this experiment.
Additionally you will be paid £[#] for your performance in Part [1, 2 or 3]. Your total earnings today are £[#]. Please wait at your desk until the experimenter comes to you.
Thank you for your participation!

A2 Experimental Instructions for the Pilot Experiments

Instructions

Welcome! You are about to participate in an experiment on decision making.

In this experiment, you will be asked to complete five tasks. Each task is independent from each other. The instructions for each task will be shown on your screens before the start of each task. Please read the instructions on your screens carefully and take notes if necessary as you will not be able to go back to read them again once a task starts.

After you have completed all of the five tasks, you will be asked to fill a survey questionnaire that will end the experiment. At the end of the experiment, one task will be randomly chosen by the computer. You will receive detailed feedback about your performance score and earnings in the selected task. Your final payment will be determined according to your earnings in the selected task. Any one task is equally likely to be selected at the end of the experiment, so you have to perform as well as you can in each task.

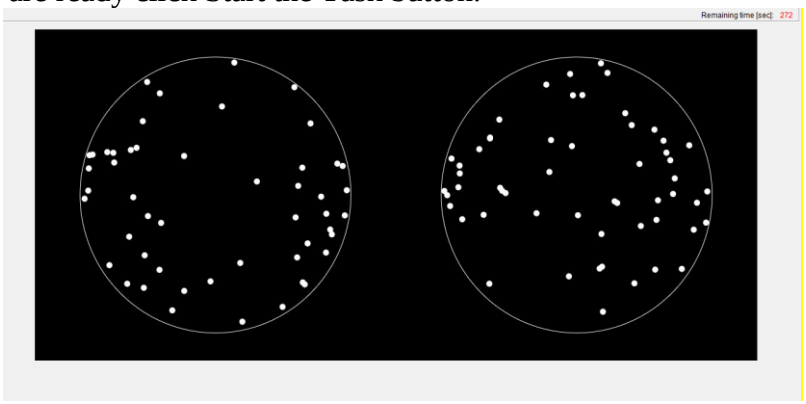
Throughout the experiment you must not communicate with other participants. The use of any electronic devices is strictly prohibited. If you break these rules, you will be excluded from the experiment without receiving any payment and be disqualified from future experiments with us.

If you have any questions before the start of the experiment, please raise your hand and someone will come to your desk to answer it.

Circle task

In this task, you will be asked to pick a circle with more dots in a series of circle pairs. The figure below shows an example of a circle pair. One of the circles contains 50 dots and the other one 55 dots. Each circle pair will appear on your screens for 1 second only. After the circle pair disappears from your screen, you will be asked to judge whether the right or the left circle contained more dots. You have to indicate your judgement by clicking on the "Left" or "Right" button. When you click the button of your choice, you will move to the next pair of circles. The task consists of three sets of 20 circle pairs and you will be notified when you finish one set of 20 circle pairs on your screens. Your performance score in this task will be the total number of correctly judged circles. You will earn £0.25 per each correct answer.

If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready click Start the Task button.

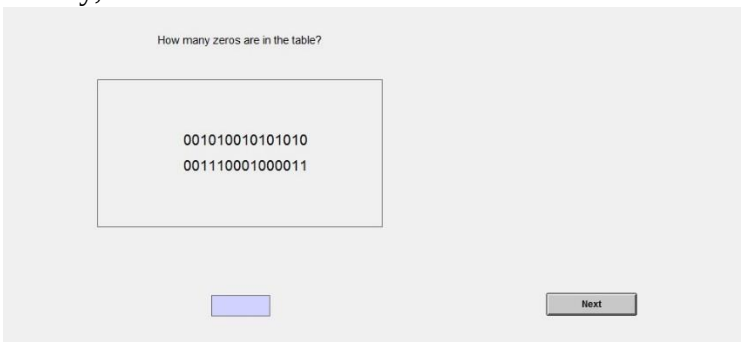


Counting Zeros task

In this task, you will be asked to count zeros in a series of tables. The figure below shows the work screen you will work on for this task. You will enter the number of zeros into the box below the table. After you have entered the number, you can click the NEXT button. No matter if the answer is correct or not, a new table will be generated. Your performance score in this task will be the number of correctly solved tables at the end of the 300 seconds. You will earn £0.40

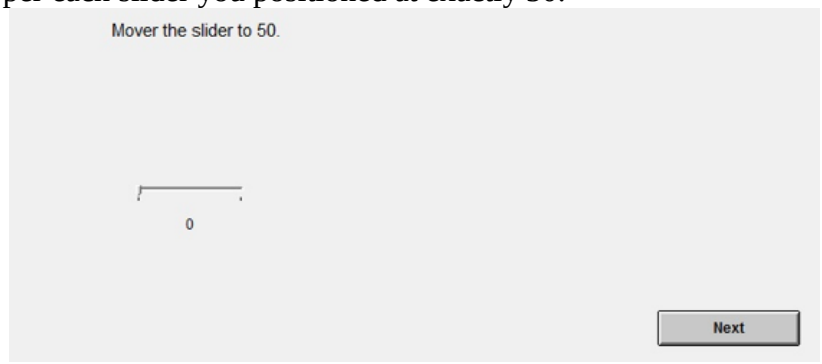
per each table you solved correctly. If you enter a wrong number for a table, you will earn nothing for that table.

If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready, click Start the Task button.



Slider task

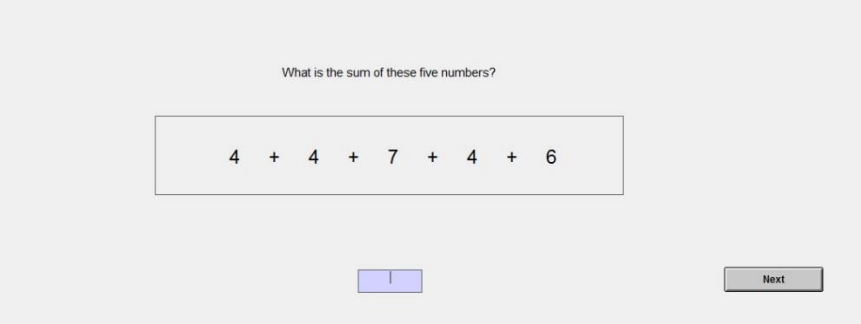
In this task, you will be asked to position a series of sliders. The figure below shows the work screen you will work on for this task. Each slider is initially positioned at 0 and can be moved as far as 100. Each slider has a number below showing its current position. You can use the mouse or the arrow keys on the keyboard in any way you like to move each slider. You can readjust the position of each slider as many times as you wish. After you have positioned a slider, you can click the NEXT button and a new slider will be generated. Your performance score in this task will be the number of sliders positioned at exactly 50 at the end of the 300 seconds. You will earn £0.10 per each slider you positioned at exactly 50.



Number Adding task [Pilot 1]

In this task, you will be asked to add five single digit numbers in a series of tables. The figure below shows the work screen you will work on for this task. You will enter the answer into the box below the table. After you have entered the answer, you can click the NEXT button. No matter if the answer is correct or not, a new table will be generated. Your performance score in this task will be the number of correct answers at the end of the 300 seconds. You will earn £0.40 per each correct answer. If you enter a wrong answer, you will earn nothing for that table.

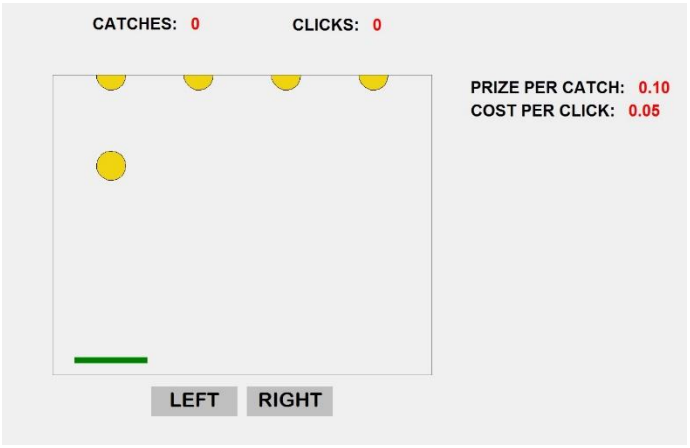
If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready, click Start the Task button.



Ball Catching Task [Pilot 1]

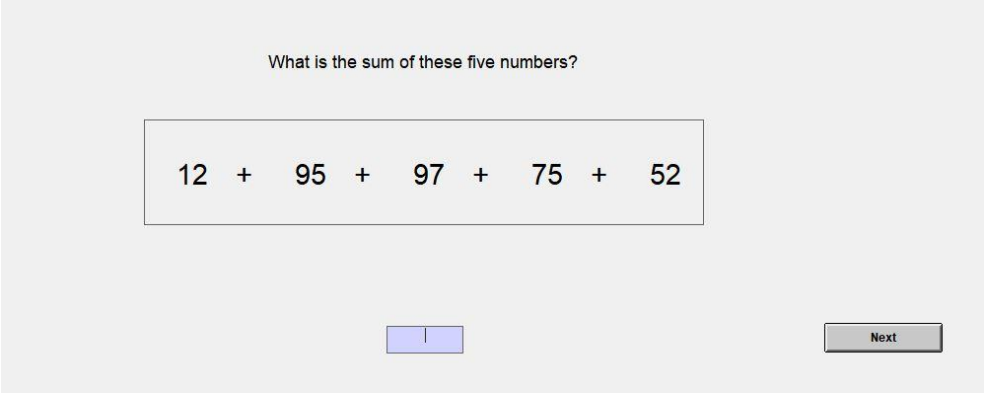
In this task, you will be asked to catch balls. There will be a box in the middle of the task screen like the one shown below. Once you click on the "Start" button, balls will fall randomly from the top of the task box. You can move the tray at the bottom of the task box to catch the balls by using the mouse to click on the LEFT or RIGHT buttons. To catch a ball, your tray must be below the ball before it touches the bottom of the tray. The number of balls you caught so far (displayed as CATCHES) and the number of clicks you made so far (CLICKS) are shown right above the task box. You will receive a prize of £0.10 for each ball you catch and incur a cost of £0.05 for each mouse click you make. Your performance score will be (CATCHES)*£0.10 - (CLICKS)*£0.05 at the end of the 300 seconds.

If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready click Start the Task button.



Number Adding task [Pilot 2]

In this task, you will be asked to add five single digit numbers in a series of tables. The figure below shows the work screen you will work on for this task. You will enter the answer into the box below the table. After you have entered the answer, you can click the NEXT button. No matter if the answer is correct or not, a new table will be generated. Your performance score in this task will be the number of correct answers at the end of the 300 seconds. You will earn £0.50 per each correct answer. If you enter a wrong answer, you will earn nothing for that task. If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready, click Start the Task button.



Anagram Task [Pilot 2]

In this task, you will be asked to solve a series of anagrams consisting of 5 to 7 letters. An anagram is a word formed by reordering the letters of another word. The figure below shows the work screen you will work on for this task. You will enter the anagram into the box below the letters. After you have entered the anagram, you can click the NEXT button. No matter if the answer is correct or not, new letters will be generated. Your performance score in this task will be the number of correct anagrams at the end of the 300 seconds. You will earn £0.50 per each correct anagram. If you enter a wrong anagram, you will earn nothing for that task.
If you have any question, please raise your hand and someone will come to your desk to answer it. If you are ready click Start the Task button.

[Eliciting Preferences for Relative Performance Pay for Each task]

Everyone has now finished all the tasks. Before proceeding to the payment stage, you have a chance to increase your earnings. You can decide whether you want to compete in each of the five tasks against one randomly selected participant.
After you have indicated your choice to compete or not for every task, the computer will randomly choose one task. If you have chosen to compete in that task, we will compare your performance score on that task to the score of one randomly selected other participant. If your score is higher than the score of the other participant, we will add £2 to your final earnings. If your score is lower than the score of the other participant, we will subtract £2 from your final earnings. If you chose not to compete, your final earnings will not be affected and you will be paid according to your performance score only on the selected task.
Please note that your choice and earnings will not affect the earnings of the other participants.
At the end of the session, you will receive detailed information about your own, the randomly selected participant's score, your choice to compete or not and your earnings resulting from your scores and choices on the selected task.

If you have any question please raise your hand and someone will come to your desk to answer it. If you are ready, please indicate your choices on the next screen.
Please indicate whether you choose to COMPETE or NOT COMPETE for each of the following tasks

Circle Task	Compete □	□ Not compete
Ball Catching Task	Compete □	□ Not compete
Slider Task	Compete □	□ Not compete
Counting Zeros Task	Compete □	□ Not compete
Number Adding Task	Compete □	□ Not compete

Questionnaire 1

Please answer the following questions trying to be as accurate as you can. Your answers are anonymous and will not be linked to your identity.

What is your gender? □ Male □ Female

What is your age? _____

What is your nationality?

□ British □ Other (please specify) _____

If you are a student, what is your subject area?

□ Humanities □ Natural Sciences □ Engineering □ Medical Sciences

□ Economics □ Business □ Politics □ Law

□ Other (please specify) _____

On a scale of 1 to 7, how willing are you to take risks in general?

1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □

Not at all willing Very willing

On a scale of 1 to 7, how confident are you as a person?

1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □

Not at all confident Very confident

On a scale of 1 to 7, how competitive are you as a person?

1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 □

Not at all competitive Very competitive

On a scale of 1 to 7, how difficult was each task that you completed?

Circle Task	1 □	2 □	3 □	4 □	5 □	6 □	7 □
Ball Catching	1 □	2 □	3 □	4 □	5 □	6 □	7 □
Number Adding	1 □	2 □	3 □	4 □	5 □	6 □	7 □
Counting Zeros	1 □	2 □	3 □	4 □	5 □	6 □	7 □
Slider	1 □	2 □	3 □	4 □	5 □	6 □	7 □
Anagram [Pilot 2]	1 □	2 □	3 □	4 □	5 □	6 □	7 □

Questionnaire 2 [Pilot 2]

On a scale of 1 to 7, please choose whether the cognitive skill described is important for high performance in each task that you completed:

Attention: Holding attention while completing the task

Circle Task	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □	Very Important □ Can't Judge
Number Adding	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □	Very Important □ Can't Judge
Counting Zeros	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □	Very Important □ Can't Judge
Slider	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □	Very Important □ Can't Judge
Anagram	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □	Very Important □ Can't Judge

Working Memory: Ability to learn information and use that information for the current activity

Circle Task	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Number Adding	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Counting Zeros	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Slider	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Anagram	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge

Visual Perception: Ability to see and interpret the visual information

Circle Task	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Number Adding	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Counting Zeros	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Slider	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Anagram	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge

Cognitive Flexibility: Being able to consider several solutions or plans, not only the first one that comes to mind

Circle Task	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Number Adding	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Counting Zeros	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Slider	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Anagram	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge

Numeracy Skills: Ability to reason and to apply simple numerical concepts

Circle Task	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Number Adding	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Counting Zeros	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Slider	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge
Anagram	Not at all Important	1 □	2 □	3 □	4 □	5 □	6 □	7 □ Very Important	□ Can't Judge

Please make a judgement whether the performance in the two tasks are positively related (people who do well in one task also are more likely to do well in the other task), negatively related (people who do well in one task are more likely to do worse in the other task), or not related. If you can't make a judgement please choose the last option.

□ The two are not related □ Positively related □ Negatively related □ I can't say/ I don't know

Circle and Slider task	Slider and Counting Zeros Task
Circle and Number adding task	Slider and Anagram Task
Circle and Anagram task	Number Adding and Anagram Task
Circle and Counting Zeros task	Number Adding and Counting Zeros Task
Slider and Number Adding task	Counting Zeros and Anagram task

B Pilot Results

Table B1: Pilot Results on Five Real Effort Tasks

Pilot 1 (N=18)	Number adding *(single-digit)*	Slider	Circle	Counting zeros	Ball catching
			Mean (SD)		
Score	40 (8.23)	46 (14.45)	41 (5.26)	32 (6.10)	123 (20.88)
Difficulty	3.00 (1.32)	3.11 (1.23)	5.16 (1.20)	3.50 (0.93)	3.05 (1.11)
Compete Choice	50%	29%	28%	33%	55%
			Spearman correlation coefficients in performance		
Slider	0.128				
Circle	0.377	-0.207			
Count Zeros	0.403*	0.417*	-0.033		
Ball Catching	0.139	-0.038	0.479**	0.355	
			Spearman correlation coefficients in perceptions of difficulty		
Slider	-0.029				
Circle	0.369	0.256			
Count Zeros	0.378	0.456**	0.312		
Ball Catching	0.194	0.254	0.274	0.212	
			Spearman correlation coefficients in choosing competitive pay		
Slider	-0.167				
Circle	0.091	0.433*			
Count Zeros	0.349	0.433*	0.150		
Ball Catching	0.528**	-0.167	0.349	0.349	
Pilot 2 (N=24)	Number adding *(double-digit)*	Slider	Circle	Counting zeros	Anagram
			Mean (SD)		
Score	10 (3.66)	42 (9.75)	41 (4.22)	29 (6.20)	21 (7.07)
Difficulty	3.58 (1.48)	2.66 (1.17)	4.08 (1.38)	2.91 (1.31)	4 (1.32)
Compete Choice	67%	71%	38%	51%	42%
			Spearman correlation coefficients in performance		
Slider	0.267				
Circle	0.133	0.199			
Count Zeros	0.087	0.356*	0.142		
Anagram	0.103	-0.038	0.545***	0.155	
			Spearman correlation coefficients in perceptions of difficulty		
Slider	0.143				
Circle	-0.025	0.018			
Count Zeros	0.295	0.632***	0.004		
Anagram	0.605***	0.028	-0.023	0.125	
			Spearman correlation coefficients in choosing competitive pay		
Slider	0.907***				
Circle	0.183	0.307			
Count Zeros	0.768***	0.697***	0.367*		
Anagram	0.598***	0.542***	0.567***	0.608***	
			Correlations with Number Adding Task in Cognitive Skills		
Attention		0.314	0.562***	0.674***	0.396**
Working Memory		0.049	0.380*	0.510**	0.521***
Cognitive Flexibility		0.225	0.331	0.622***	0.165
Visual Perception		0.314	0.211	0.235	0.659***
Numeracy		0.164	0.075	0.334	0.141
			Relatedness Scores with Number Adding: % choosing either not related or don't know.		
Slider	Circle	Counting Zeros	Anagram		
65%	50%	20%	40%		

** 10%, ** 5%, *** 1% significance levels.*

C Additional Tables and Figures for the Main Experiment

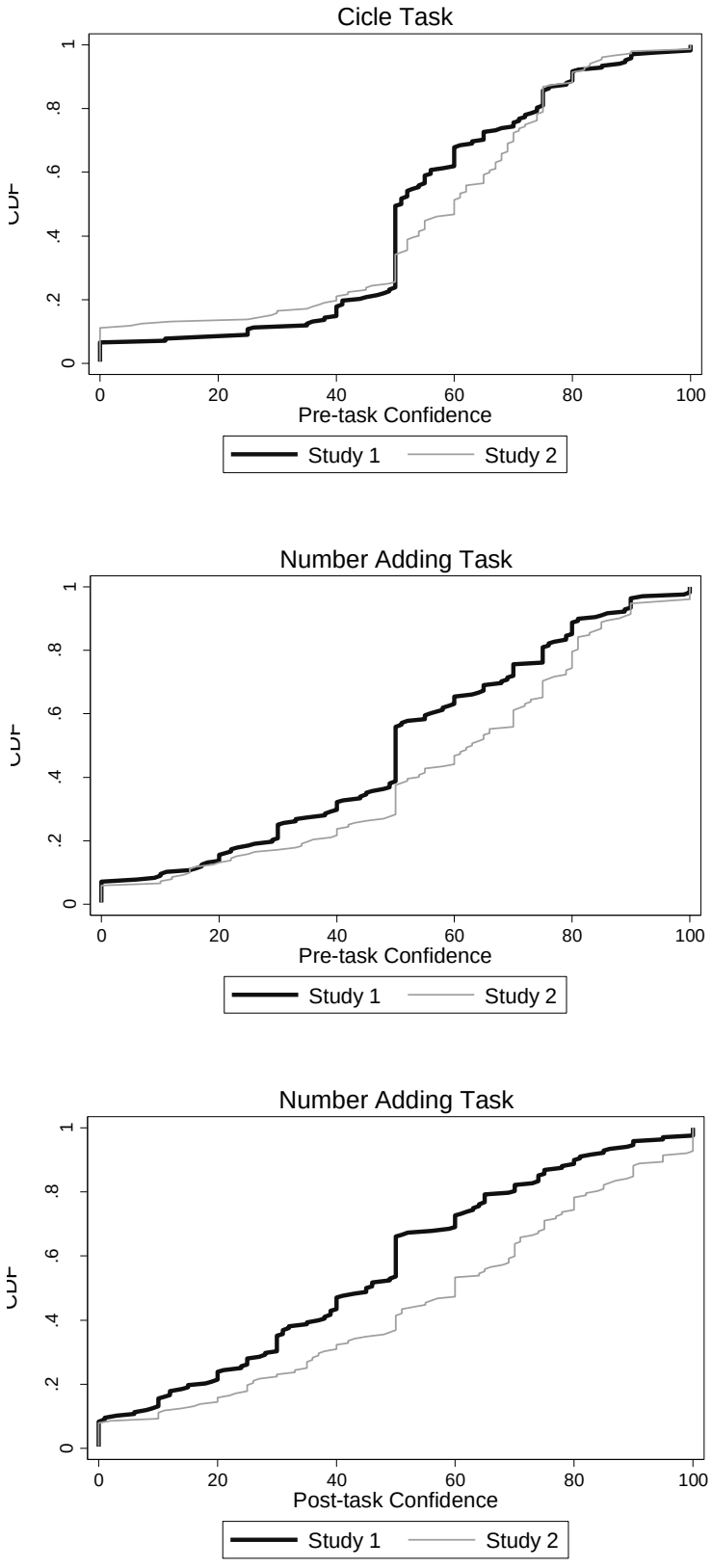


Figure C1: Elicited Confidence in Study 1 and Study 2

Table C1a: Robustness Tests for Feedback Spillover Effects (Study 1)

Dependent Variable	Pre-task Confidence (Number)		Post-task Confidence (Number)		PR-Equivalent	
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Symmetrically censored least squares regression with the Table 1 controls						
Top (vs bottom)	15.317***	14.603***	6.723	7.671	0.206*	0.306*
	(4.771)	(5.079)	(6.775)	(7.470)	(0.121)	(0.159)
Top (vs no feedback)	5.48	5.080	3.126	2.733	0.090	0.076
	(4.616)	(4.153)	(5.250)	(4.986)	(0.085)	(0.085)
Bottom (vs. no feedback)	-6.356	-6.314	-1.719	-4.324	-0.102	-0.159
	(4.980)	(5.080)	(5.950)	(5.492)	(0.103)	(0.102)
(b) Tobit Regression controlling for second order circle task score polynomial						
Top (vs bottom)	14.266***	13.615***	8.765	9.170	0.232***	0.205**
	(4.795)	(4.222)	(7.637)	(7.487)	(0.084)	(0.094)
Top (vs no feedback)	2.217	2.147	3.100	3.045	0.092	0.089
	(4.022)	(3.511)	(4.228)	(4.601)	(0.087)	(0.088)
Bottom (vs. no feedback)	-6.374	-5.934	-1.044	-1.268	-0.077	-0.059
	(5.617)	(4.705)	(5.596)	(5.181)	(0.090)	(0.089)
(c) Tobit Regression controlling for third order circle task score polynomial						
Top (vs bottom)	15.445***	14.894***	8.949	5.554	0.243***	0.225**
	(5.213)	(5.105)	(7.675)	(6.645)	(0.086)	(0.107)
Top (vs no feedback)	3.077	2.882	4.561	4.221	0.094	0.092
	(5.007)	(4.522)	(4.218)	(4.376)	(0.101)	(0.102)
Bottom (vs. no feedback)	-8.640	-8.074	-0.795	-1.294	-0.127	-0.115
	(6.477)	(5.570)	(4.896)	(4.705)	(0.106)	(0.102)
(d) Tobit regressions with number adding score fixed effects						
Top (vs bottom)	20.354***	20.501***	9.843	10.150	0.271***	0.259***
	(5.172)	(6.006)	(6.790)	(6.686)	(0.029)	(0.029)
Top (vs no feedback)	3.921	3.486	1.998	1.968	0.074***	0.081***
	(4.143)	(3.508)	(4.789)	(5.094)	(0.020)	(0.022)
Bottom (vs. no feedback)	-8.911	-9.096*	-3.039	-3.176	-0.113***	-0.104***
	(6.450)	(4.846)	(5.090)	(4.855)	(0.022)	(0.024)
Demographics & psych. Controls	N	Y	N	Y	N	Y

Note: Standard errors clustered at session level are in parentheses. All regressions include controls of circle task confidence, number adding task score, and circle task score (specification varying across the panels b, c and d). Number of variables are the same as reported in Table 1.

Table C1b: Robustness Tests for Feedback Spillover Effects (Study 2)

Dependent Variable	Pre-task Confidence (Number)		Post-task Confidence (Number)		PR-Equivalent	
(a) Symmetrically censored least squares with Table 1 controls						
Top (vs bottom)	6.127	5.727	15.488***	13.959**	0.118	0.111
	(4.444)	(4.465)	(5.587)	(5.751)	(0.072)	(0.069)
Top (vs no feedback)	0.037	-0.610	1.623	0.631	0.076	0.093
	(4.161)	(4.271)	(4.840)	(4.635)	(0.070)	(0.065)
Bottom (vs. no feedback)	-8.676*	-8.050	-11.120*	-8.495	-0.062	-0.027
	(5.127)	(5.143)	(5.763)	(5.897)	(0.082)	(0.074)
(b) Tobit Regression controlling for second order circle task score polynomial						
Top (vs bottom)	6.553	4.822	20.050**	17.462**	0.162**	0.113
	(5.120)	(5.314)	(7.971)	(7.051)	(0.066)	(0.073)
Top (vs no feedback)	1.310	0.133	3.331	1.957	0.098***	0.093***
	(5.471)	(5.885)	(5.253)	(5.193)	(0.035)	(0.033)
Bottom (vs. no feedback)	-5.871	-5.754	-8.073	-6.115	-0.053	-0.040
	(6.536)	(6.456)	(5.461)	(4.701)	(0.069)	(0.061)
(c) Tobit Regression controlling for third order circle task score polynomial						
Top (vs bottom)	3.667	2.271	20.086**	18.226**	0.205**	0.162
	(7.046)	(7.256)	(7.893)	(7.862)	(0.078)	(0.102)
Top (vs no feedback)	3.077	2.882	4.561	4.221	0.094	0.092
	(5.007)	(4.522)	(4.218)	(4.376)	(0.101)	(0.102)
Bottom (vs. no feedback)	-8.640	-8.074	-0.795	-1.294	-0.127	-0.115
	(6.477)	(5.570)	(4.896)	(4.705)	(0.106)	(0.102)
(d) Tobit regressions with number adding score fixed effects						
Top (vs bottom)	5.516	5.831	20.691***	19.769***	0.184	0.155
	(5.220)	(4.432)	(6.332)	(6.263)	(0.118)	(0.138)
Top (vs no feedback)	-2.160	-2.259	6.948*	5.239	0.063	0.070
	(5.500)	(5.829)	(4.135)	(4.089)	(0.065)	(0.064)
Bottom (vs. no feedback)	-2.164	-1.882	-9.126	-7.834	-0.082	-0.088
	(7.262)	(7.559)	(6.570)	(6.802)	(0.075)	(0.075)
Demographics & psych. Controls	N	Y	N	Y	N	Y

Note: Standard errors clustered at session level are in parentheses. All regressions include controls of circle task confidence, number adding task score, and circle task score (specification varying across the panels b, c and d). Number of variables are the same as reported in Table 1.

Table C2: Feedback spillover effects on preferences to participate in experiments with relative performance pay

Dependent Variable is the answer to the question: "We will be holding different types experiments in the future. If you had a choice, what kind of experiments would you like to participate in?" 1 – Strong preference for relative performance pay experiments, 5 – strong aversion to relative performance pay experiments. (Study 2)				
	(1)	(2)	(3)	(4)
Top (vs no	-0.233	-0.091	-0.198	-0.083
feedback)	(0.251)	(0.296)	(0.242)	(0.296)
Bottom (vs. no	0.526**	0.486**	0.462**	0.427*
feedback)	(0.176)	(0.210)	(0.192)	(0.222)
Pre-task Confidence			-0.002	-0.002
(Circle)			(0.005)	(0.004)
Pre-task Confidence			-0.010**	-0.007*
(Number)			(0.003)	(0.004)
Constant	3.153***	5.334***	3.883***	5.803***
	(0.165)	(0.776)	(0.258)	(0.850)
Top vs. Bottom p-value	*0.022*	*0.093*	*0.046*	*0.140*
Demographics & psych. controls	N	Y	N	Y
Obs.	152	152	152	152

Standard errors clustered at session level are in parentheses. Demographics include gender and age. Psychological measures include general risk-taking, confidence, and competitiveness. Top vs. Bottom p-values are from post-estimation Wald tests.