Southern Methodist University

SMU Scholar

Summer 8-6-2024

Experiments on Employer-Employee Relationships

Hang Chen
hangc@smu.edu

Recommended Citation

EXPERIMENTS ON EMPLOYER–EMPLOYEE RELATIONSHIPS

Approved by:

_____

Dr. Tim Salmon
Professor of Economics

_____

Dr. Santanu Roy
Professor of Economics

_____

Dr. Daniel Millimet
Professor of Economics

_____

Dr. Krista Saral
Associate Professor of Economics

EXPERIMENTS ON EMPLOYER–EMPLOYEE RELATIONSHIPS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Economics

by

Hang Chen

B.S., Mathematics, Sun Yat-sen University
M.S., Economics, University of Wisconsin-Madison
M.A., Economics, Southern Methodist University

August 6th, 2024

ACKNOWLEDGMENTS

Hang Chen
B.S., Mathematics, Sun Yat-sen University
M.S., Economics, University of Wisconsin-Madison
M.A., Economics, Southern Methodist University

Experiments on Employer-Employee Relationships

Advisor: Dr. Tim Salmon

Doctor of Philosophy conferred August 6th, 2024

Dissertation completed June 18th, 2024

Employer-employee relationships exist commonly in workplaces and are important for people to understand the interactions between employers and employees and also their own behavior. This dissertation presents studies of different aspects of employer-employee relationships using laboratory experiments.

Chapter 1 studies how the problematic nature of many middle management positions drive middle managers to interact with their employees in a negative way in workplaces. There are more and more attention paid to negative work relationships that middle managers treat their employees badly. Middle managers are blamed for their behavior, but they are also under pressure from upper management to extract high effort from their employees, but given few incentives to do so. The question we ask in this study is whether this situation, by its nature, is one that can create middle managers' negative behavior to employees. We conduct an experiment using a modified version of a gift-exchange game in which a manager has the opportunity to provide feedback, positive or negative, to a worker as their main means of encouraging effort. The managers will face varying levels of pressure nominally from upper management for eliciting effort from the worker to see how pressure of this sort impacts the relationship dynamics. We also investigate whether a reward bonus with a small amount of monetary incentive may help diminish negativity in the workplace. Our findings support the notion that indeed much of the "toxicity" in many work environments could stem from the problematic nature of many middle management roles and perhaps even a

small increase in the ability of the middle managers to reward employees could help out.

Chapter 2 focuses on the potential peer effects of employees' performance feedback and performance. In a group with one middle manager and multiple employees, employees can not only receive their own performance feedback from the middle manager, but also be able to observe their peers' performance feedback. The potential peer effects of performance feedback may exist and alter employees' behavior from what we have found in the essential employer-employee relationships in workplaces. It will be more complicated if they can also check peer workers' performance as the peer effects of performance can possibly show up. I implement an experiment extended from the one that characterizes the basic middle manager-employee relationship to investigate these new interactions and potential peer effects among employees. My current data suggests that there is no overall treatment effect of peer workers' performance feedback or performance on employees' effort choices. A positive correlation between coworkers' efforts and employees' efforts is found and indicates the peer effects of performance can potentially exist, which echos the results from a number of studies about peer effects.

Chapter 3 examines if females and minorities have reduced willingness to participate in tasks with more subjective judgment. It is commonly observed that women and ethnic/racial minorities are often underrepresented in some professions while they obtain better (though not always equal) representation in others. Prior research has identified many different reasons why underrepresentation might exist. Here I examine a new possible contributor to this problem, which is that people who expect discriminatory treatment, i.e. women and ethnic/racial minorities, may have preferences for jobs where they expect to be evaluated on objective rather than subjective criteria perhaps due to a hope that this might decrease their experience of discrimination. I investigate the existence of these preferences using an economic experiment in which I allow subjects to choose between jobs where their earnings depend explicitly on the evaluation of another person or where their earnings are determined by a purely objective process. I include treatments in which the human evaluator does versus does not have access to the demographic information of the worker. My data provides

suggestive but not significant results to support my hypotheses that correlations between subjects' avoidance to subjective judgment and their genders, ethnicities, or demographic information being accessible to the judge exist. Correlations between subjects' changes of avoidance to judgment and their genders or ethnicities are found insignificant as well. I discuss ideas for what this suggests about the possibility of such preferences and what might be required to observe them, if they exist.

TABLE OF CONTENTS

LIST OF FIGURES

Dedicated to My New Long Spring

CHAPTER 1

The Plight of the Middle Manager

1.1 Introduction

The Great Resignation has led to a substantial increase in the attention paid to working environments. This is because many of those participating in the "Great Resignation" have said they have quit their jobs because of what they describe as very bad working conditions in their previous jobs (Sull et al., 2022).[1] This exodus of workers is certainly a sign that the workers are suffering,[2] but turnover of this size is also bad for employers. A survey by Predictive Index in November 2021 shows that around 20% of the total workforce quit in the prior 6 months with the cost of $11,372 for each employee on average. The cost of turnover due to workplace culture exceeded $223 billion from 2015 to 2019 (SHRM, 2019). These facts make it clear that understanding the dynamics of these work environments is important to improve working conditions for employees and prevent employers from facing such mass resignations.

A standard structure to most companies is that there are at least three layers of individuals. At the top, there are the owners or upper managers who typically retain most of the decision-making power over earnings of all employees. At the bottom of the organization

---

[1]Clear data on this is difficult to obtain but one important form of a negative work environment would be workplace bullying, which is a collective expression that includes various forms of mistreatment and hostile behavior in the workplace setting (Fox and Stallworth, 2005). A meta-analysis conducted by Nielsen et al. (2010) shows that workplace bullying happens to a shockingly high 14.8% of the workforce.

[2]Victims of workplace bullying take 7 additional days off work per annum on average, and self-rate their level of productivity 7% lower than do their non-bullied counterparts (Hoel and Copper, 2001). Abused employees are at an increased risk of developing depression, post-traumatic stress disorder, alcohol abuse, stress disorder, hypertension, or committing suicide (Lutgen-Sandvik et al., 2007).

is the workers who are necessary for producing whatever product or service the company provides. In the middle are the middle managers who are the ones who are immediate supervisors of the workers. It is these managers who are usually tasked with keeping the workers productive, making sure they show up on time and so on. On a day-to-day basis, they typically do not have the authority inside the organization to provide much in the way of incentives to the employees. In most organizations, even in the long term the power to change the pay level or provide substantial bonuses is held by the upper management leaving the middle managers without many tools to use in encouraging employees to do their jobs well.

So how do these middle managers try to get their workers to work hard? One tool that is always left to them is verbal persuasion. That is, managers might try to convince their workers to work hard by praising them for good effort or castigating them for providing lesser effort. There is a great deal of debate about which of these forms of encouragement might work best. A particularly notable example of this type of behavior can be found in how high school or college sports coaches deal with their athletes. These are cases where the coach can't actually pay their players,[3] yet they face intense pressure to get them to perform. Some coaches will go to the extreme in their negative treatment of players including the famed case of Bobby Knight, a former basketball coach at Indiana University, throwing chairs at his players he thought were underperforming (Wertheim, 2015). Mike Krzyzewski, the longtime coach of the Duke basketball team, is however famous for believing that the best way to motivate athletes is through positive encouragement (Krzyzewski, 2011). There is a long literature on which is the better coaching style such as the study by Mertens et al. (2018), with the general finding that positive motivation should work best which might be considered ironic given how many coaches become famous for their more negative style.

While the sports context provides a number of colorful examples of this issue, there is also a great deal of literature on the issue from the management world. Farnsworth et al.

---

[3]A notable example of a university of attempting to do otherwise is seen when the SMU football program violated the NCAA rules as it recruited athletes by using "under the table" payments during late 1970s and early 1980s, and received "death penalty" in 1987 (Eric Dodds, 2015).

(2020) claim that positive recognition for good performance will result in performance improvement, while the improvement from negative feedback will likely be only minimum and temporary. In contrast, Hirsch (2018) provides several suggestive approaches to delivering negative feedback as it can be effective. Much of this literature discusses these issues but does not quite provide definitive proof. As an attempt to provide that proof, there have been many experiments conducted to study the effect of positive or negative feedback. In Peeters and Vorsatz's (2013) work, immaterial rewards and sanctions both increase cooperation in a voluntary contribution experiment, and rewards are more effective than sanctions, while Dugar (2013) finds that disapproval points generate higher contribution than the approval points in a public good game. Kandul and Lanz (2021) suggest that positive and negative out-group feedback have opposite directions of effects on distinct types of in-group contributors in a public good experiment. While these papers show that either might work, what is of great importance is which type of feedback might managers prefer to use, which do they think works and whether the type of feedback they choose is related to the amount of pressure they themselves face.

In order to examine these questions, we will use a modified version of the gift-exchange game to represent the core relationship between a middle manager and a worker. In the classic gift-exchange game (Fehr et al., 1993), a manager sets a wage for a worker and the worker chooses how hard to work. It is key to this game that the manager's earnings are increasing in the effort of the worker and that the worker's pay comes out of the earnings of the manager. In many organizations, a middle manager might set the pay for a worker upon hiring, within some bounds, but the amount paid to the worker does not actually diminish the earnings of the middle manager. On the other hand, it is still the case that the earnings of the middle manager will be related to the work product of the workers he or she supervises. This disconnect between the standard gift-exchange game and how wage offers are usually made is important for this situation as the fact that the wages of the worker are costless to the manager may significantly diminish the effect of high wages being seen as a "gift." Under this paradigm, the top wage may be perceived as "standard" or "expected" while any wage less than the top one available may be seen as poor treatment from the employee's per-

spective. Given that this relationship is perhaps more standard for wage formation than the usual gift-exchange game it is important to examine this structure to see if the results of the gift-exchange game translate into this paradigm or if the change to the nature of the wages changes how wages are perceived leading to a different relationship between wages and effort.

A standard finding in the classical gift exchange game is that effort is positively related to the wage offered even though that effort is not incentivized. We of course may find a different result in our base environment. To further examine how managers can affect employee effort we will be allowing our manager and worker to interact over several rounds, but as in most workplaces, the wage will be fixed after the initial choice. Any further impact the manager can have on effort levels will come through their ability to provide feedback to employees. We will allow positive and negative feedback to be available simultaneously in a manner that follows previous laboratory experiments adopting feedback of approval or disapproval points as non-monetary incentives (Masclet et al., 2003; Peeters and Vorsatz, 2013; Lefebvre and Stenger, 2020). Many of the prior papers only allow a single form of feedback, positive or negative, but Cooper and Lightle (2016) allow for both in an examination of how feedback options change reciprocity. Their feedback though is from the worker to the manager, which is the opposite direction of our focus. While these studies have demonstrated that feedback of this nature can impact behavior, it is not clear whether positive or negative feedback from a manager can increase effort levels of worker.

The managers in our environment will be modeled to be facing differential pressure from upper management. This will take the form of an effort threshold which they must induce their employee to surpass in order to receive earnings in a round. We vary that minimum to see how it impacts feedback.

In anticipation of observing some negativity on the part of the managers, we also wanted to examine whether allowing managers the ability to make small financial rewards could solve the problem. We do this by allowing employers to allocate a bonus to their employee that is of an amount able to offset a small amount of the employee's effort cost. We also

model the award language off of how many employer's do allocate small awards to employees. While one might imagine that it is obvious that this approach should work, there is prior evidence that small monetary awards may crowd out and reciprocal behavior and result in lower overall effort (Gneezy and Rustichini, 2000; Titmuss, 1970; Mellstrom and Johannesson, 2008).

Our findings show that despite the fact that setting high wages is costless, many managers choose not to set wages at the highest level. We do not, however, see evidence that our modified gift exchange game eliminates the reciprocal elements from wage formation as we find that effort is still positively related to the wage. Our managers often employ negative feedback for underperforming employees and those who face higher pressure exhibit negative behavior more frequently and intensely. Managers' feedback has no overall treatment effect on employees' effort choices, which implies that sending feedback to employees, as a tool for middle managers to apply, is not able to encourage increases of effort in general. We observe a correlation between negative feedback employees receive in one period and higher efforts they exert in the next period, which indicates that negative feedback can potentially motivate low performing employees to increased efforts. While many prior studies have suggested that positive reinforcement might motivate employees, we find, perhaps sadly, that middle managers seem only to consider positive feedback for well performing employees, and there is no correlation between the positive feedback and employees' higher effort choices. Given that managers' feedback is endogenous and we mostly observe negative feedback to low effort levels and positive feedback to high effort levels, we have to be careful about how broadly we can interpret and apply these results. Our findings provide suggestive explanations on why many managers may choose to use negative feedback and how their situation may generate some amount of toxicity in a work environment. An encouraging finding is that our reward bonus is able to have a modest impact on increasing effort while significantly reducing the amount of negative feedback sent. Thus even a very small flexibility on the part of the managers may be effective at countering problematic work environments to everyone's benefit.

1.2 Experiment

1.2.1 Design

The base interaction in our experiment is based on the standard gift-exchange game but we implement an important change. At the beginning of the experiment, subjects are assigned in pairs with one member of the pair assigned as the manager and the other as the employee. This pairing and these roles will be constant throughout the experiment. In the first round of our experiment, the manager will set the wage of the employee from the range [0,20] ECUs.[4] The employee can then choose an effort level on the range [0,10]. The payoff to the employee will be twice the offered wage less the effort choice. The clear incentives there are that for any wage offered, the employee maximizes earnings with an effort choice of 0. This part of the game is exactly the same as the standard gift-exchange game.

The manager's earnings are where our game differs. In many workplaces, while a middle manager may be allowed to set the wage of a new worker from small pre-defined range upon hiring them, that wage does not come out of their own earnings or salary. The effort of the worker, however, will determine the productivity of that manager's unit and could therefore impact whether that manager retains their job and if they do, how much they earn. We will represent this by assigning a manager a minimum effort level they must elicit from their employee that is required for the manager to receive any earnings. If the effort level of the worker is at or above that level the manager receives earnings of 20 added with doubled effort. The payoffs of the paired manager and the employee show as follow:

$$\Pi_M = \begin{cases} 0, & \text{if } e < \overline{e} \\ 20 + 2e, & \text{if } e \geq \overline{e} \end{cases}$$

$$\Pi_E = 2w - e$$

Notice that a significant difference between this and the classic formulation is that the worker's wage does not enter into the manager's earnings. This can change the expectation on the part of the worker. A high wage may no longer be seen as a gift but rather it may be simply expected that the manager should offer the top wage given that it costs them nothing. Further if a manager chooses to offer a lesser wage, even a high one, the worker could well see that as a negative choice which they then reciprocate negatively instead of positively. All of these incentives are made transparent to workers and managers except that the exact value of the effort level cutoff required for the manager to receive earnings is not communicated to the worker. When making their effort choices, workers are given a payoff calculation tool on their screen to show them for any level of effort they might choose, what their own payoff will be and then what the payoff level for the manager might be should the effort level cross the required threshold or not.

Each manager was given a minimum effort cutoff of either 2 or 5. These effort requirements are randomized between middle managers with both being equally likely. The assigned threshold remained fixed for a manager for the duration of the experiment. The manager is informed of their threshold but the employee is not.

The manager and worker will interact over 3 parts of the experiment for a total of 15 periods. The wage for the worker is set in the first period and remains constant for all subsequent rounds. This reflects the standard wage formation structure in many workplaces which is that while a manager may have some impact on the wages of an employee when hired, those wages will remain the same over a long period of time. During that time period,

7

the manager still has to extract effort from the employee and so our experiment deals with that interval of time.[5]

In part 1 of the experiment (Baseline treatment), the manager is allowed to choose a wage for the employee in the first period, and this wage will be fixed in the next 4 periods. So in period 1, the manager chooses a wage and the worker then chooses effort as described above with no feedback. In periods 2-5, the manager has no active role, the worker is reminded of the wage and then the worker chooses effort.

In part 2 of the experiment (Feedback treatment), each round will proceed as before with the worker being reminded of their wage and then choosing effort. After that effort level is chosen, the manager will see that effort level and will be able to send feedback in the form of happy or sad face emojis. The manager can send multiple emojis, up to 10 sad face emojis or up to 10 happy face emojis. They can only send happy or sad emojis, they cannot send a combination of both. The worker will see this feedback prior to the next round.

In part 3 of the experiment (Feedback & Bonus treatment), we implement our reward bonus by adding on one more element to the previous part. In addition to the feedback, a manager can now opt to send the employee a bonus of 2 ECUs on the same screen. Employees and middle managers are told that the bonus is meant to be an award to recognize the employee's valuable effort contribution in a period. This monetary incentive is a modest amount which does compensate the employee for 2 units of effort. Thus if a worker believes the employee will provide the bonus for 2 units of effort, now there could be an equilibrium with positive levels of effort. This will be below the threshold required in the high-pressure case for the manager and so we deem this as a nominal bonus, not one capable of fully incentivizing effort.

---

[5]One might be concerned that the wage of the employee could eventually rise to reward them for strong effort. In some cases that is certainly true. Whether that employee receives that wage is typically determined by upper management though, not the middle manager and many workplaces will have tenure-based wage systems rather than merit-based systems. Again, these are the environments our design represents. For workplaces where managers can reward employees, then obviously our design will not capture those situations.

One challenge to identifying treatment effects in this design is the fact that effort could change over time due to issues unrelated to the treatments. To account for this we also conduct sessions in which all 15 periods use the baseline structure, meaning that neither feedback nor bonus are allowed. These control sessions allow us to use a standard difference-in-difference method to test the treatment effects of Feedback and Feedback & Bonus. One might also be concerned that we have used only a single ordering of our Feedback and Feedback & Bonus treatments. This is intentional. The question we wish to address with the latter treatment is whether it can increase effort and/or decrease the use of negative feedback from the middle treatment. It would make little sense to test that in a reverse order.

1.2.2 Implementation

We conducted 11 experimental sessions at the Laboratory for Research in Experimental Economics (LREE) at Southern Methodist University and there were 92 participants in total. These 11 sessions included 7 treatment sessions with 64 subjects and 4 baseline sessions with 28 subjects. The experiment was programmed in z-Tree (Fischbacher, 2007) and subjects were recruited among pre-registered LREE students using an hroot based recruitment system (Bock et al., 2014). Each subject participated in only one session and was grouped and assigned a role randomly.

We read aloud a script to explain laboratory rules and experiments general introduction after subjects were seated, and detailed instruction of a treatment before each part of the experiment. Subjects were also guided through sample decision-making screens before each treatment.

In order to guarantee anonymity, subjects were not asked to reveal their actual names or other information during the experiment, though we did have to collect and record names and payments for accounting purposes in the end. As mentioned, all monetary amounts in the experiment were denominated in ECUs with the rate of 100 ECUs = \$4.00. Each session lasted around 1 hour, with average earnings of \$24.3 per subject, including a \$10 show-up

fee.

1.2.3 Hypotheses

We now present the hypotheses our experiment was designed to test. Our first hypothesis relates to the core gift-exchange game relationship in our model. In the standard gift-exchange game it is observed that high wages generally lead to higher levels of effort as those high wages may be perceived as a "gift" by the worker and so they reciprocate the generosity by choosing high effort. In our version, the high wage is costless to the manager which may undermine its value as a gift and expectations could change to simply assuming that it is obvious that the manager should offer the high wage. If so, then the standard gift-exchange relationship may not hold. We can use our data to test if it does. This leads to our first hypothesis which we will phrase as testing whether the standard relationship holds.

Hypothesis 1.1. Employees choose higher effort in response to higher wage.

Our next hypotheses concern the general treatment effects in our design. As discussed above, there are those who believe the positive feedback may induce employees to work hard while others believe that negative feedback is more effective. The problem is that either could well backfire. If a manager provides positive feedback to an employee, that feedback might make the employee feel valued and encourage them to work hard to receive more. The positive feedback could also be seen as a sign that the employee is already working hard enough which could lead to backsliding on effort or at least not increasing it. If one is already receiving positive feedback, why work harder? For negative feedback, an employee receiving it could decide to work harder so that they no longer experience that negativity. The negative feedback may also upset an employee and discourage them from further effort. Given these possibilities it is entirely unclear whether either will be effective in encouraging effort and whether there will be any overall impact on effort from managers having the possibility to provide feedback.[6] We will state the hypotheses in a positive frame due to a prevailing

---

[6]It is possible that feedback like this could work a bit differently if a manager has multiple employees with the ability to see feedback that others receive. We do not discount this possibility rather we wish to start out examining this simpler environment as a first test of the issue and leave the more complicated

belief that feedback of various sorts might improve effort.

Hypothesis 1.2. Employees choose higher effort during the Feedback treatment than under the Baseline treatment.

Our next treatment effect focuses on whether the Feedback & Bonus treatment can further increase effort.

Hypothesis 1.3. Employees choose higher effort during the Feedback & Bonus treatment than during the Baseline treatment.

If the Feedback treatment has an overall positive or negative impact on effort choices, that is not enough to tell us what type of feedback might be accomplishing the shift in effort. We therefore need to test more specifically how employees respond to positive or negative feedback. One important issue with these tests is that unlike the treatments which are exogenously imposed, the specific nature of feedback one receives is endogenous. This will be important in examining these issues and trying to draw inference on the impact of different forms of feedback on effort. For now we will state two related hypotheses suggesting that both types of feedback could improve effort choices, but as we discussed above, there are credible reasons to think we could find an alternate result.

Hypothesis 1.4a. Positive feedback will be associated with increases in effort choices.
Hypothesis 1.4b. Negative feedback will be associated with increases in effort choices.

The results of testing these hypotheses will help us understand how employee's choices differ based on the treatments and the choices by the managers. In analyzing them, we will need to take into account the circumstances under which managers provide each type of feedback to try to understand this issue more fully. Our core questions are really centered on the managers' behavior in regard to when and how severely they use negative feedback.

multi-employee dynamics for future research.

11

We want to understand whether middle managers' use of negative feedback is more frequent and more intense when they are under higher pressure. Of course the primary way that we represent pressure in this experiment is based on the effort requirement the manager faces, so we will investigate how feedback varies with that requirement. An important related issue is how feedback might change depending on when effort is above versus below the threshold requirement for the manager. Our hypotheses here are that managers facing pressure are more likely to use negative feedback and that they are most likely to use feedback on employees performing under the threshold. This leads to two hypotheses with two different ways of stating/testing the issue.

Hypothesis 1.5a. Middle managers choose negative feedback more frequently when their effort requirement is high.

Hypothesis 1.5b. Middle managers choose negative feedback with increased intensity when their effort requirement is high.

Hypothesis 1.6a. Managers send negative feedback more often when the effort level of their employee is below the effort requirement.

Hypothesis 1.6b. Managers send negative feedback with higher intensity when the effort level of their employee is below the effort requirement.

When the bonus is available, it seems reasonable that it could serve as either a substitute or a complement to the feedback. Awarding the bonus could replace positive feedback as this is feedback that actually should matter to the employee. Withholding the bonus could replace negative feedback. This is important as it is possible for the availability of the bonus to eliminate the need for a manager to use negative feedback. Consequently we wish to examine what happens to the managers' choices of feedback during the phase when the bonus is also allowed.

Hypothesis 1.7a. Managers choose negative feedback less often when the bonus is available.

Hypothesis 1.7b. Managers choose negative feedback with less intensity when the bonus is available.

An important consideration here is the fact that the bonus could impact feedback in two ways. First, as described above, the bonus could serve as a substitute for feedback. The bonus also has the possibility of having a direction impact on the effort chosen by the employee and if it increases effort, that increase in effort might also diminish any use of negative feedback by the manager. In testing this hypothesis, we will need to be careful to identify the impact of the availability of a bonus on the use of feedback separate from the impact on the use of feedback from any change in effort due to the availability of the bonus.

1.3 Results

1.3.1 Summary Statistics

We first present the summary statistics of middle managers' and employees' choices to give a brief view of our data. The employees' summary statistics are shown in Table 1.1, where we lay out the average wage, average feedback[7] and frequency of bonus the employees receive and average effort they exert in both types of sessions in the three different parts. Employees choose efforts on average well above the theoretical prediction of 0 in all segments. The average effort levels in parts 1 and 2, Baseline and Feedback, appear quite similar but there is an indication that effort may rise during the second treatment stage, Feedback & Bonus. The choices of middle managers are summarized by treatment in Table 1.2. This table shows that average wage that they offer, average feedback (positive and negative) in part 2 and part 3, and frequency of bonus in part 3 they choose are listed in the table and divided by effort requirements, together with the average effort they receive from employees. From this we see that middle managers with higher effort requirement choose higher wages for employees and receive higher efforts from them. We also see an indication that managers facing the higher effort requirement choose to send more negative feedback than those with the lower effort requirement.

---

[7]We define the value of feedback as from -10 to 10 which corresponds to 10 unhappy faces to 10 happy faces.

	All	Part 1	Part 2	Part 3
(Baseline, 14 Subjects)				
Average Wage	15.86	15.86	15.86	15.86
Average Effort	4.17	4.16	4.23	4.13
(Treatment, 32 Subjects)				
Average Wage	14.28	14.28	14.28	14.28
Average Effort	4.44	4.13	4.17	5.03
Average Feedback	2.00	-	1.00	2.99
Frequency of Bonus	98	-	-	98

Table 1.1: Employees' Summary Statistics

	All	Requirement=2	Requirement=5
Number of Subjects	32	16	16
Average Wage	14.28	13.43	15.13
Average Effort	4.44	3.73	5.16
Average Feedback	2.00	2.34	1.66
Average Positive Feedback	4.85	4.30	5.52
Average Negative Feedback	-4.58	-3.50	-5.40
Frequency of Bonus	98	48	50

*Part 2 & 3; **Part 3

Table 1.2: Managers' Summary Statistics

1.3.2 Employees' Behavior

From the employees' summary statistics, we already find that they exert positive effort, instead of choosing zero effort as theory suggests, which indicates that there may exist a reciprocal effect in the effort choice based on the wage offered by the managers. For a clearer indication of such a relationship, we can examine Figure 1.1 which shows the relationship between wage offered and the corresponding average effort choice in the first period and then the same between the wages and the average effort employees chose over all periods from our sessions. In both cases, we see a positive correlation between employees' efforts and wages.

14
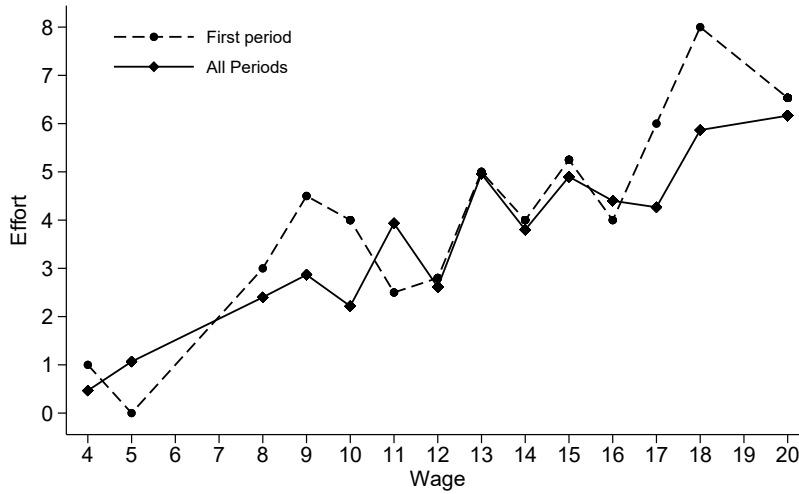
Figure 1.1: Average effort to wage in the first period and all periods.

While Figure 1.1 provides a suggestive relationship, we conduct a set of regressions to provide formal tests of our first hypothesis. The first regression in Table 1.3 is an OLS regression of effort on wage restricted to period 1. The second regression is a random effects panel regression using data from all periods where the dependent variable is effort, and independent variables are wage, feedback, feedback & bonus, part 2 and part 3. Standard errors are clustered by subjects. Here feedback and feedback & bonus are treatment variables indicating when the treatments are in effect, while part 2 and part 3 are phase variables indicating which part of the experiment participants are at, as we also have baseline sessions that have no treatment. Recall that we have the set of baseline sessions where all phases are run under the baseline treatment so that we can use this to identify any trends in behavior simply due to time spent in the interactions. When we test for our treatment effects, this allows us to conduct a difference-in-differences regression to get clear identification of those effects. For now, we can note that both regressions confirm that there is a significant positive correlation between effort and wage confirming that the standard gift-exchange effect translates into our modified gift exchange game.

| | (1) | (2) |
	Effort	Effort
Wage	0.339***	0.351***
	(0.0766)	(0.0558)
Feedback		0.0506
		(0.230)
Feedback & Bonus		1.013***
		(0.373)
Part 2		0.0170
		(0.158)
Part 3		-0.0830
		(0.289)
Constant	-0.220	-1.046
	(1.181)	(0.725)
Observations	46	690
R-squared	0.308	
Number of Subjects	46	46

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.3: (1) OLS regression of effort on wage restricted to the first period. (2) Random effects regression of effort on wage, treatment variables and phase variables.

Result 1.1. Employees choose higher effort in response to higher wage.

Our next several hypotheses examine whether employees respond to the treatments. Figure 1.2 presents the average efforts in all 15 periods in our experiment, with one line for the sessions in which the treatments were implemented and another line for the sessions with 15 periods of the baseline interaction. It is the difference between these lines that will represent any treatment effects. We see virtually no difference between the treatment and baseline sessions in periods 6-10 which suggests that the Feedback treatment may lead to no aggregate impact on effort. There appears to be a gap between the lines in periods 11-15, suggesting a treatment effect may exist for the Bonus & Feedback treatment.
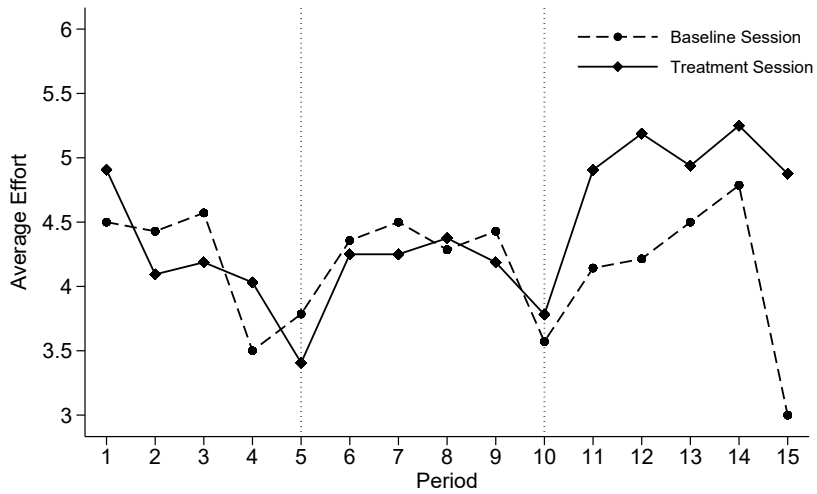
Figure 1.2: Average effort in all periods, from treatment sessions and baseline sessions.

The second regression in Table 1.3 allows us to test our hypotheses regarding these treatment effects, corresponding to hypotheses 1.2 and 1.3. We find no significant impact on effort from the Feedback treatment alone but we do find a positive and significant impact from the Feedback & Bonus treatment.

Result 1.2. There is no significant difference in effort due to the Feedback treatment.

Result 1.3. Employees choose higher efforts in due to the Feedback & Bonus treatment.

These results tell us that the existence of feedback on average has no effect, but that should not be taken as clear evidence that individual instances of feedback do not influence effort. As described above, feedback may encourage and discourage effort and this overall result could be found due to those two effects existing and canceling out in the aggregate. To examine whether feedback can have any influence on effort choices, we have to examine the data in a different way.

For this analysis it is important to understand that while the existence of the treatment or not was exogenous, giving us clear identification, the existence of specific feedback from a manager to a worker is endogenous. Thus our analysis of this issue will be looking at correlations between the two as we do not have random assignment of feedback that would be necessary to make stronger causal claims. Figure 1.3 shows how employees respond to negative feedback. It provides histograms of employees' changes of efforts from period t-1 to period t when they receive negative feedback at period t-1 (t starts from period 7) divided into one histogram for when employee effort in t-1 is below the manager's threshold and the other when their effort is above. One thing to note is that negative feedback is given frequently in both situations. The distribution of effort changes has a mean of 1.8 when effort is below the threshold and -0.1 when it is above. Figure 1.4 shows the same information but for positive feedback. The initial point to note from this is that positive feedback is virtually never chosen by managers when effort choices are below their threshold. Thus we can make no inference regarding how well such feedback might work. Positive feedback is given quite often when effort is above the threshold and the mean of the distribution of effort changes is -0.3. The indication is that positive feedback does not seem to be increasing or decreasing effort choices significantly under any specific pattern.
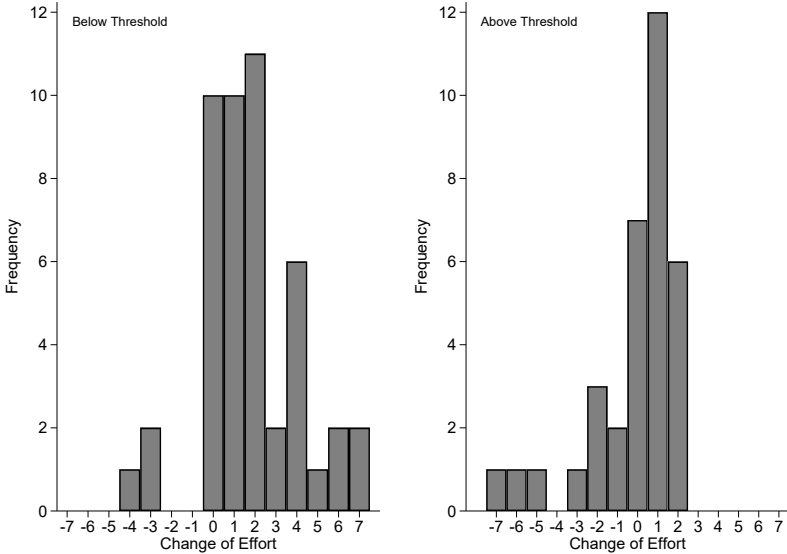
Figure 1.3: Employees' changes of effort after receiving negative feedback.
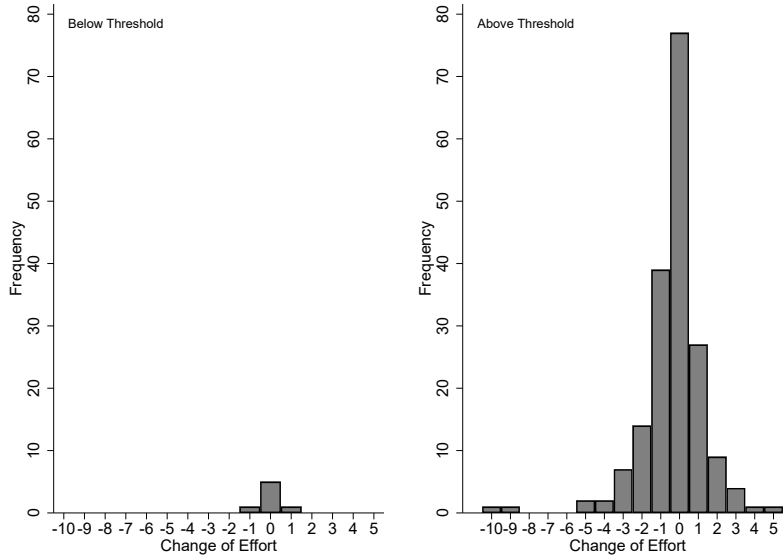
Figure 1.4: Employees' changes of effort after receiving positive feedback.

A statistical look at this relationship is shown in Table 1.4 where we provide panel regressions of employees' change of effort between periods t and t-1 as function of the feedback they received in period t-1. As before these are random effects regressions with the standard errors clustered at the subject level. Dependent variables of both regressions are change of effort between t and t-1. Independent variables in the first regression are wage, feedback & bonus, lagged negative feedback which is an indicator variable for whether the prior feedback was negative or not and lagged positive feedback indicating whether the prior feedback was positive or not. For those feedback variables, the omitted case is no feedback. As for the second regression, the independent variables are wage, bonus, lagged value of negative feedback and lagged value of positive feedback where we use the count of the sad or happy faces to measure the intensity of the feedback. The first regression tells us that there is a significant correlation between negative feedback and increase of effort from employees, and the second regression shows that lower value of negative feedback is correlated to increase of effort, while higher value of positive feedback is linked to decrease of effort. These suggest that very intense negative feedback can be useful in increasing effort while very intense positive feedback is actually associated with lowering effort. For these cases, it is important to understand where these impacts are occurring. When effort is below the threshold, man-

agers give very negative feedback and the indication is that it can be successful in increasing effort. When effort is above the threshold, managers choose a mix of positive and negative effort. It is not clear that negative feedback still helps nor that positive feedback helps. This might be because when employee effort is low, they likely recognize that and respond to negative reinforcement because they realize that it might be reasonable to consider their effort too low. For higher levels of effort, the negative feedback may lose its ability to push effort farther as employees are no longer convinced that their effort is too low even with substantial negative feedback. When an employee receives positive feedback, they seem to take the implication that their effort levels are considered acceptable and so there is no need to improve them. If feedback is too positive, employees may indeed backslide on effort as their inference might be that they are working harder than they need to.

	(1) Change of Effort	(2) Change of Effort
Wage	0.0159	0.0661**
	(0.0205)	(0.0291)
Negative Feedback (t-1)	0.965***	
	(0.351)	
Positive Feedback (t-1)	-0.430	
	(0.316)	
Feedback & Bonus	0.534***	0.629***
	(0.134)	(0.156)
Value of Negative Feedback (t-1)		-0.231***
		(0.0415)
Value of Positive Feedback (t-1)		-0.130***
		(0.0415)
Constant	-0.438	-1.114**
	(0.375)	(0.482)
Observations	288	288
Number of Subjects	32	32

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.4: (1) Random effects regressions of effort change on wage, lagged negative/positive feedback and feedback & bonus. (2) Random effects regressions of effort change on wage, lagged value of negative/positive feedback and feedback & bonus.

Result 1.4a. There is no significant difference of efforts if the employees receive the positive feedback or not.

Result 1.4b. There is a significant positive correlation between negative feedback and employees' changes of effort.

Overall the data on the choices of employees shows that the standard gift-exchange effect does emerge in our modified game and that the Feedback & Bonus treatment is effective in increasing effort. Negative feedback seems useful in shifting effort levels higher for employees with low efforts, while positive feedback is not able to increase employees' efforts even more.

1.3.3 Managers' Behavior

An initial curiosity in regard to manager behavior concerns how they chose wages. Given that the wages are costless, there is no clear theoretical prediction about what they might choose as they are theoretically indifferent. If managers believe that higher wages will lead to higher efforts, then we should expect that all managers should choose the highest wage level. The summary statistics of middle managers shows that they actually chose average wages of around 14. This average comes from finding that around 33% of the middle managers choose the highest wage while 50% of them choose wages that are less than 15. Clearly not all managers are taking advantage of their opportunity to costlessly set high wages. There are many possibilities for why this might be the case, but our experiment was not designed to examine these issues directly so we will take no inference about why they might have done this.

Our main interest in the managers focuses on their choice of what sort of feedback to send to employees and how they allocated the bonus when it was available. As we illustrated when presenting hypotheses 1.5 and 1.6, we would like to investigate the middle managers' preferences to the type of feedback and its intensity in two situations. We want to know if middle managers give negative feedback more frequently and more intensively when they are notified that they have high effort thresholds or they receive efforts that are below their effort requirements.

Average feedback chosen by middle managers having high or low effort thresholds in part 2 and 3 from treatment sessions are presented in Figure 1.5. Middle managers choose lower average feedback when they have high effort thresholds, and their average feedback is close to negative in part 2 of the experiment, when the bonus is not yet available.

We also have Figure 1.6 showing the similar middle managers' average feedback choices, and the only difference is that they are categorized by the efforts being above or below the middle managers' effort requirements. Here we see that middle managers choose lower value of average feedback when they receive efforts that are below their effort thresholds, as might be expected.
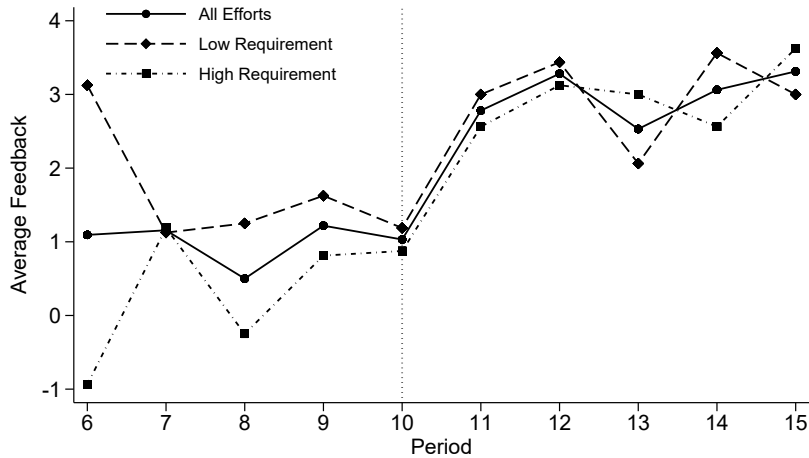
Figure 1.5: Average feedback, divided by low and high effort requirements.
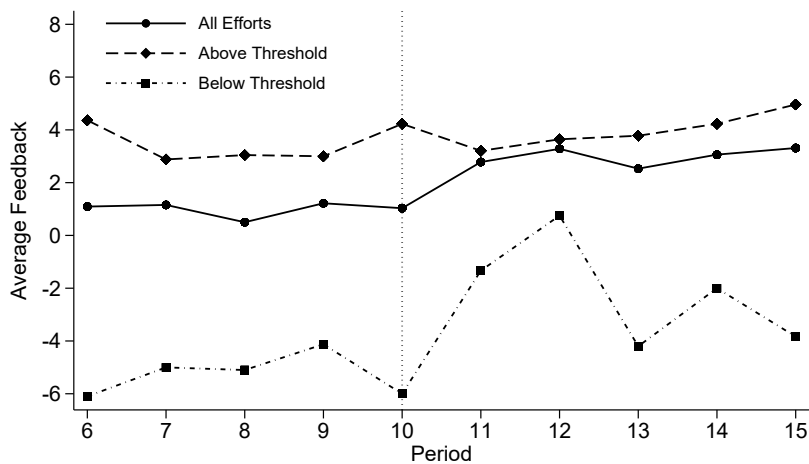
Figure 1.6: Average feedback, divided by above and below effort thresholds.

Table 1.5 gives us the results testing hypotheses 5a and 6a. There are three logit regressions with standard errors clustered by subjects and negative feedback being the dependent variable indicating feedback is negative or not. The first regression has wage, feedback & bonus, and high requirement indicating a manager's effort requirement to be high or not to be independent variables, the second regression has effort added as an independent variable, while the third one has wage, feedback & bonus, high requirement and low effort specifying effort is below the threshold or not. The second and the third regressions show strong relationships between negative feedback and high requirement or low effort, which indicates the middle managers are more likely to choose negative feedback when the effort requirement is high, or the efforts they receive are below the thresholds.

	(1) Negative Feedback	(2) Negative Feedback	(3) Negative Feedback
Wage	-0.153*	0.173	-0.00152
	(0.0833)	(0.121)	(0.0918)
Effort		-1.467***	
		(0.395)	
Feedback & Bonus	-1.242***	-0.444	-0.880**
	(0.310)	(0.520)	(0.361)
High Requirement	0.823	3.436***	0.119
	(0.683)	(1.328)	(0.794)
Low Effort			4.375***
			(1.088)
Constant	0.892	0.607	-2.276
	(1.293)	(1.587)	(1.615)
Observations	320	320	320
Number of SubjID	32	32	32

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.5: (1) Logit regression of negative feedback on wage, feedback & bonus and high requirement. (2) Logit regressions of negative feedback on wage, effort, feedback & bonus and high requirement. (3) Logit regressions of negative feedback on wage, feedback & bonus, high requirement and low effort.

Result 1.5a. Middle managers choose negative feedback more often when the effort requirement is high.

Result 1.6a. Middle managers choose negative feedback more often when the effort is below the requirement.

The results of three random effects regressions are listed in Table 1.6. The dependent variables in these regressions are value of negative feedback. The first regression's independent variables are wage, feedback & bonus and high requirement, the second regression has one more independent variable effort, while the third one's independent variables are wage, feedback & bonus, high requirement and low effort. Standard errors are clustered by subjects as well. We still find significant positive correlation between the intensity of negative feedback and requirement being high or effort being below the threshold from the second and the third regressions.

	(1) Value of Negative Feedback	(2) Value of Negative Feedback	(3) Value of Negative Feedback
Wage	0.193**	0.0276	0.124
	(0.0815)	(0.0949)	(0.0789)
Effort		0.942***	
		(0.207)	
Feedback & Bonus	1.825***	1.296*	1.479**
	(0.584)	(0.673)	(0.648)
High Requirement	-1.920**	-3.003***	-1.279
	(0.858)	(0.534)	(0.795)
Low Effort			-2.634***
			(0.599)
Constant	-6.469***	-6.011***	-4.322***
	(1.349)	(1.171)	(1.356)
Observations	88	88	88
Number of Subjects	24	24	24

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.6: (1) Random effects regression of value of negative feedback on wage, feedback & bonus and high requirement. (2) Random effects regressions of value of negative feedback on wage, effort, feedback & bonus and high requirement. (3) Random effects regressions of value of negative feedback on wage, feedback & bonus, high requirement and low effort.

Result 1.5b. Middle managers choose negative feedback more intensively when the effort requirement is high.

Result 1.6b. Middle managers choose negative feedback more intensively when the effort is below the requirement.

As for our last hypothesis, we try to find out if the reward bonus can drive the middle managers to choose negative feedback less frequently and intensively when it is available. In

Figure 1.5 and Figure 1.6, we observe that middle managers choose better feedback (less negative feedback and higher feedback value) in part 3 of the experiment, in which the reward bonus is available, especially middle managers with high effort requirements and receiving efforts below the thresholds. The first regression in Table 1.5 and the first regression in Table 1.6 show that there are significant correlations between feedback & bonus and negative feedback or value of negative feedback, which means that the middle managers are less likely to use negative feedback, or apply negative feedback with less intensity when the reward bonus is available. Here we still need to identify how the availability of reward bonus affects the middle managers' use of negative feedback, as we notice that employees increase their efforts when the reward bonus is available from the second regression in Table 1.3, and higher efforts lead to middle managers' use of negative feedback to be less frequent and intense from the second regressions in Table 1.5 and Table 1.6. The same regressions in Table 1.5 and Table 1.6 also include feedback & bonus as independent variables, and there is negative relationship between feedback & bonus and negative feedback, and positive relationship between feedback & bonus and value of negative feedback, which indicates that when the reward bonus is allowed to assign, the middle managers' improved behavior of applying negative feedback is not only indirectly caused by increased efforts, but also directly driven by the presence of the treatment.

Result 1.7a. Middle managers choose less negative feedback when the reward bonus is available.

Result 1.7b. Middle managers choose negative feedback less frequently when the reward bonus is available.

Middle managers in our experiment do not choose wages at highest level for their employees. When there is short term pressure that they receive efforts that are below their effort thresholds, or there is long term pressure that they find that they are assigned high effort requirements, they decide to send negative feedback more often and more severely to employees. As for reward bonus, it helps middle managers apply less negative feedback and with less intensity.

1.4 Conclusion

In this paper we attempt to examine some important issues regarding why middle managers treat their employees badly and what might drive these negative behavior. Our primary interest was in how the plight of middle managers may contribute to these problems which refers to the fact that upper management in a firm usually places substantial pressure on middle managers to motivate their workers while providing remarkably few tools to these managers to do so. We design an experiment modifying the standard gift-exchange game model to address the middle manager-employee relationship and study their interactions. The data from our experiment shows that middle managers use negative feedback almost exclusively to low performing workers, and only once the efforts are generally acceptable, they begin to consider providing positive feedback. We simulate upper management putting stress on middle managers by implementing varied minimum effort thresholds the employees have to surpass in order for the middle managers to receive earnings and find that managers facing higher pressure will be tempted to rely on negative feedback to their employees more likely and severely. There is no causal treatment effect of feedback on employees' efforts generally, but negative feedback appears to be effective in raising underperforming employees' efforts, as there is a positive correlation found in our data between employees' negative feedback and changes of efforts. Positive feedback seems to be unable to increase well performing employees' efforts further since no correlation between positive feedback and employees' changes of efforts exists. The correlation between managers' negative feedback and employees' efforts helps explain why there lack an overall effect from the feedback, as employees increase efforts when receiving more negative feedback and reduce efforts after having less negative feedback. The changes of efforts cancel out in general. These results are important as they demonstrate how a very common management structure used in firms can be a potentially strong contributor to poor quality working environments in which managers regularly choose to use negative feedback as a means of motivating workers. While the actual feedback used in our experiment was quite mild in nature, it is quite reasonable to expect that this mildness could turn into what many people might call "toxic" in a more realistic setting.

Our results also show that there is some hope in regard to a simple way to potentially diminish the managers' use of negative feedback. This feedback is used perhaps primarily because managers have no other tools. We investigate whether allowing the managers to allocate very small incentives could alleviate the problem and find that indeed this is the case. When we allow our managers to allocate relatively small amounts of money to their employees, this mechanism replaces the negative feedback and generates better work effort. This is a potential win for employees who face less negative feedback, middle managers who feel less of a need to engage in that behavior and to the firm which would reap the reward of better performing employees who may be less likely to quit their jobs.

In the process of performing the analysis above, our paper provides a useful contribution to the voluminous literature on the gift-exchange game. In the standard and much researched version of that game, the manager offers a worker a wage which is costly to the manager. This costly offer has the possibility of establishing a reciprocal relationship as it demonstrates that the manager is giving something up for their worker which may lead to the worker feeling indebted to the manager for the gift and therefore opt to repay that gift by working hard. In many, perhaps most, employment relationships in firms, when a middle manager offers an employee a job that manager may be able to have some control over the wage to be offered, often from a pre-specified range decided upon by upper management, with the key point that higher wages do not cost the middle manager anything. They do not have to sacrifice or give up much to make a worker a more generous offer. It seems reasonable to expect that this change to the game could impact how workers see wage offers and possibly substantially change the results from what is commonly observed in the gift-exchange game. In our modified version of this game, we actually find that the standard result still holds. Employees still seem to view high wages offers as "gifts" and respond in kind with higher effort. This is an important finding when considering whether the standard gift-exchange result might apply outside of the lab.

There are some important future directions for this work to go in order to solidify these findings and to expand them to other contexts. In our environment, a single manager in-

teracts with a single employee. It is reasonable to think that the impact of feedback can change when a manager controls multiple employees and the feedback is publicly viewable. Negative feedback applied to one employee may motivate others to work hard to not receive that feedback while positive feedback given to one might motivate others to work harder to receive the same. These effects would be similar to the ones we find above but seeing the feedback others receive informs employees on the counterfactual possibilities of what might happen to them with different effort levels and that can change how effective feedback is due to the externalities. The implementation of a small bonus may also be more effective when a group dynamic is considered as winning the bonus may be seen as more of a prize to compete for if it is limited to a single member of a group or team. Such a structure could allow an award to be effective for even lesser nominal amounts of money as the pride or prestige from the scarcity of the award could serve as the motivation to receive it.

Overall the results of our study help us understand why many workplaces may develop a dynamic that leads to unhappy employees and overly stressed middle managers. Our results also demonstrate that there may be some simple ways to change these environments for the benefit of all involved. Future work is of course necessary to verify the best corrective measures in different work environments but the fact that we resolved the problem relatively easily in this situation provides strong hope that other corrective measures should be available for many workplaces currently operating under the problematic model we identify in this study.

CHAPTER 2

Peer Effects of Middle Managers' Feedback

2.1 Introduction

In Chen and Salmon (2023), middle managers are shown that they can treat their employees badly when they are under the pressure from upper management's high requirement and employees' low performance. The reason that middle managers present such negative behavior to their employees is that they are not authorized to provide any monetary incentives to encourage employees to work harder. In other words, it is the institutional framework that all these people work at that causes middle managers' negative treatment to employees and therefore following problems such as toxicity in workplaces. Employees, especially those with lower performance, seem to respond to middle managers' negative feedback by raising efforts. These findings about the middle manager-employee relationship match a number of observations in the real world. For example, college sports teams' coaches treat athletes badly in order to push them to perform better as they are not authorized to assign any monetary incentives. Chain restaurants' managers, who cannot raise wages or provide monetary bonus to employees in daily work, are found to use negative verbal feedback to make them work harder. And these negative treatments are useful in increasing players and workers' performance, especially when they exert relative low efforts. However, there still remains an important difference between the framework in the study and field cases: there are multiple middle manager-employee relationships in a group, i.e. a team includes one middle manager leading several employees, and the peer effects among employees can exist and potentially alter the employees' behavior.

Peer effect appears as long as individuals do not exist in isolation but are embedded within networks of relationships, such as families, coworkers, neighbors, friendships or socio-economic groups. It is simply the indirect effect of the behavior of a social reference group on individual outcomes (Welteke, 2015) due to conformism, social pressure, rivalry, social learning and distributional preferences (Villeval, 2020). Peer effects can exist in various places and scenarios. For instance, coworkers significantly shape individuals' fertility behavior as individuals follow the fertility decisions of their coworkers of the same age group and in the same occupation (Paola et al., 2024), and utilities can reduce energy consumption by providing feedback to customers on home electricity and natural gas usage with a focus on peer comparisons (Ayres et al., 2013).

In a group with a middle manager and multiple employees in workplaces, employees' effort behavior can be affected by the peer effects of performance feedback from the middle manager to other workers. It is possible that they present positive responses to their peer employees' negative feedback from the middle manager as they realize they need to work harder to avoid the same bad judgment, or they might simply increase their efforts knowing their peers' positive feedback in order to catch up. When it comes the observability of peer employees' efforts as an addition, their own efforts can change even further by the peer effects of performance of coworkers as they potentially increase the efforts to follow coworkers' good performance, or just respond to coworkers' bad efforts by working harder.

A great deal of researches about peer effects of performance suggest that there are positive influences if workers are able to observe their peers' efforts. Mas and Moretti (2009) find strong evidence of positive productivity spillovers from the introduction of highly productive supermarket cashier into a shift. Falk and Ichino (2006) study subjects who are asked to fill letters into envelopes in pairs or alone and find that peer effects raise productivity. However, a number of studies also show that there are no peer effects or peer effects vary. Guryan et al. (2009) use random assignment in professional golf tournaments to test for peer effects and find no evidence that playing partners' ability affects performance. One interpretation is that professional golf players' performance links mostly to their abilities instead of willingness or

efforts in a specific game. The same explanation applies to Fan et al. (2022), who examine individuals' performance on a job with quality and quantity tasks and suggest that peer pressure might not be able to promote workers' productivity, as individuals may not raise their performance regarding to quality quickly. The results of a lab experiment designed to study the role of observability for peer effects in the setting of a simple production task by Veldhuizen et al. (2018) show some workers reciprocating a high-productivity co-worker but others taking the opportunity to free ride.

In addition to the peer effects of performance, there are also potential peer effects of performance feedback among employees. It is hard for employees in workplaces to avoid receiving feedback from middle managers not only to themselves but also to their coworkers, and the peer effects of this performance feedback can possibly drive employees to change their effort behavior. There lack studies that systematically examine the peer effects of performance feedback in literature and only a few studies about peer effects in the scenario with one principal and multiple agents find that agents can read messages from the principal to other peer agents and react. For example, Zimmerman (2003) uses data on college students' grades, their SAT scores, and the SAT scores of their roommates to measure peer effects in academic outcomes. He finds that peer effects are almost always linked more strongly with verbal SAT scores as students in the middle of the SAT distribution may have somewhat worse grades if they share a room with a student who is in the bottom 15% of the verbal SAT distribution. Here the SAT scores can be considered as performance feedback but it may not have direct effects on students' college grades, since it won't update or last along students' college studies and their grades can be more affected by their daily interactions. Duflo et al. (2008) study the academic performance of students in tracking schools in Kenya and find there exists positive direct effect of high-achieving peers, but again the effects of tracking as students' performance feedback can be limited as it doesn't happen frequently. A close study is from Horton (2010), which includes employer's expectations. In his study, workers are hired to perform an image-labeling task, and evaluate the work product of other workers in some cases. There are also employer's expectations in the form of messages sent to workers about how the task should be finished. The messages of how the task should

be done can be understood as the employer's evaluations to all the workers' performance. Workers are found to increase productivities after evaluating high-output works, but some of them would ignore the employer's expectations and are affected more by coworkers' high performance.

Although previous studies show that either the peer effects of performance or the peer effects of performance feedback can exist and have impacts on workers' effort choices, it is more important to find out how employees react to both possible peer effects at the same time in workplaces, especially on the daily basis. I would like to investigate what might happen if employees' interactions are included in a group as those potential peer effects can complicate employees' effort behavior. When employees work towards a job in the same group, and they all decide the middle manager's earnings, will they react to the performance feedback as evaluations from the middle manager to themselves as before, and will they respond to the feedback from the middle manager to their fellow workers? In addition to being able to observe coworkers' feedback, employees can also check coworkers' performance - if the observability of peer employees' efforts has any effect on employees' effort choices?

I implement an experiment to find answers to these questions. The experiment is modified from the experiment characterizing the essential middle manager-employee relationship (Chen and Salmon, 2023) with only one obvious change: there are one middle manager and two employees in each group. Now the middle manager sets a wage for two employees and each employee chooses effort and then affects the middle manager's earnings independently. The middle manager will observe employees' efforts and send them feedback separately and as for the reward bonus, the middle manager can choose only one employee to assign to. This experiment therefore has two independent middle manager-employee relationships in a group and allows me to investigate the interactions between employees. Depending on the types of the experiment sessions, in certain parts of the experiments, an employee is able to check the peer employee's performance feedback from the middle manager or the peer employee's performance feedback together with efforts. I can therefore identify the potential peer effects of performance feedback and performance as employees are assigned to different

kinds of interactions and information, compared to the situation that there is one single employee in a group and no possible extra interactions.

My current data from the experiment suggests that there is no overall treatment effect of employees' own performance feedback, peer workers' performance feedback, or peer workers' performance on employees' effort choices. But a negative correlation between own performance feedback and employees' efforts, and a positive correlation between coworkers' performance and employees' efforts can be found in the data, which suggest that employees seem to increase efforts when receiving their own negative performance feedback or knowing peer employees' high performance. To briefly conclude, there are indications that employees in workplaces react to their own feedback, the peer effects of performance might exist as they follow peer employees' efforts, but there are no peer effects of performance feedback.

2.2 Experiment

2.2.1 Design

As mentioned, the experiment in this study is modified from the experiment (Chen and Salmon, 2023) with a single middle manager-employee interaction in a group. Most of the settings of the experiment keep the same while I make some necessary changes in order to study those questions raised at the beginning.

Every one middle manager and two employees form a group. The grouping and subjects' roles will be fixed along the whole experiment. The experiment includes three parts representing three treatments, and each part has five periods. In the first period of the experiment, the middle manager in each group will choose the wage to both employees from the range [0,20] ECUs.[8] As it has been found that employees' effort choices are affected by the wage set by the middle manager, here I keep the wage identical to two employees in the group to control its effect on employees' efforts. The middle manager's effort requirement later is the same to two employees' tasks for the same reason, although the tasks and the

---

[8] All monetary amounts in the model and experiment were denominated in ECUs (Experimental Currency Units) which translate into dollars at the rate of 100 ECUs = \$4.00.

corresponding earnings are independent to the middle manager, because effort requirement will decide the middle manager's feedback choice behavior, and therefore varies employees' efforts. Each employee will then choose an effort level ranging [0,10]. The payoff to the employee will be twice the wage less the effort.

The middle manager's earnings still depends on his/her own effort requirement, which represents the expectation from the upper management, and each employee must contribute an effort higher than the threshold to generate positive payoff to the manager. Here the middle manager's overall earnings comes from two employees' efforts independently. The manager receives earnings of 10 added with doubled effort from each employee, if the effort is above the requirement. The payoffs of the middle manager M and each employee E_i are shown as follow ($i \in \{1, 2\}$):

$$\Pi_M = \Pi_{M_1} + \Pi_{M_2}$$

$$where \;\; \Pi_{M_i} = \begin{cases} 0, & \text{if } e_i < \overline{e} \\ 10 + 2e_i, & \text{if } e_i \geq \overline{e} \end{cases}$$

$$\Pi_{E_i} = 2w - e_i$$

Again, the employees' wage does not come from the middle manager's own earnings. As the result, the wage here may not be considered as a costly gift from the manager and therefore employees will not show the gift-exchange effect and simply choose zero effort to maximize their own benefits. All of these incentives are made transparent to employees and the middle manager except that the exact value of the effort level threshold required for the manager to receive positive earnings is not communicated to the employees. Subjects are able to calculate their and group member's payoffs on their decision-making screens, with all these variables interact.

The middle manager is given a minimum effort threshold of either 2 or 5 with equal probability. The assigned threshold remains fixed for the manager for the duration of the experiment and is identical to both employees' tasks. The manager is informed of his/her threshold but the employees are not.

The wage will be kept constant along the course of the experiment, which means the middle manager can only set the wage in the first period of the experiment. Employees, however, will keep choosing efforts in every periods of the experiment. The interactions after the employees' choices of efforts vary in different treatments of the experiment.

In part 1 of the experiment (Baseline treatment), the middle manager chooses a wage for both employees in the first period, and this wage will be fixed in the next 4 periods. So in period 1, the manager chooses a wage and the employees then choose efforts independently as described above. In periods 2-5, the manager has no active role, the employees are reminded of the wage and then choose efforts.

In part 2 of the experiment (Feedback treatment), each round will proceed as before with the employees being reminded of their wage and then choosing efforts. After that efforts are chosen, the middle manager will see those efforts and be able to send feedback in the form of happy or sad face emojis to two employees separately. The middle manager can send multiple emojis, up to 10 sad face emojis or up to 10 happy face emojis but can only send happy or sad emojis. A combination of both cannot be sent. The employees will see their own feedback prior to the next round.

In part 3 of the experiment (Feedback & Bonus treatment), the reward bonus is added to the previous part. In addition to the feedback, the middle manager can now assign a bonus of 2 ECUs as a reward to an employee. Employees and the middle manager are told that the bonus is meant to be an award to recognize one employee's valuable effort contribution in a period. In other words, the middle manager can only choose one employee out of two to assign the reward bonus. This small amount of monetary incentive does not come from

the manager's earnings as well.

This experiment design branches off into two kinds of sessions, depending on the information employees are able to observe and receive in a group. In Peer Feedback session, employees in a group can see their peer employees' feedback from the managers as long as the feedback starts to be sent (part 2 and part 3), and in Peer Feedback & Effort session, employees can check their peer employees' feedback from the manager, together with their efforts in a period (part 2 and part 3). Considering that the Treatment session of the pervious study (Chen and Salmon, 2023) includes exact the same three parts but with only one employee in a group, it can be used and labeled as the Baseline session which has the identical Feedback and Feedback & Bonus treatment but no peers' feedback and effort observable. Table 2.1 shows different sessions and treatments that help identify the effects of peer's feedback and efforts.

	Part 1	Part 2	Part 3
Baseline Session	Baseline Treatment	Feedback Treatment	Feedback & Bonus Treatment
Peer Feedback Session	Baseline Treatment	Feedback Treatment (Peer's Feedback)	Feedback & Bonus Treatment (Peer's Feedback)
Peer Feedback & Effort Session	Baseline Treatment	Feedback Treatment (Peer's Feedback & Effort)	Feedback & Bonus Treatment (Peer's Feedback & Effort)

Table 2.1: Sessions and treatments for identification.

2.2.2 Implementation

There were 6 experimental sessions conducted at the Laboratory for Research in Experimental Economics (LREE) at Southern Methodist University and 24 participants in total. These 6 sessions included 3 Feedback sessions with 12 subjects and 3 Feedback & Effort sessions with 12 subjects. The experiment was programmed in z-Tree (Fischbacher, 2007) and subjects were recruited among pre-registered LREE students using an hroot based recruitment system (Bock et al., 2014). Each subject participated in only one session and was grouped and assigned a role randomly.

I read aloud a script to explain laboratory rules and experiments general introduction after subjects were seated, and detailed instruction of a treatment before each part of the experiment. Subjects were also guided through sample decision-making screens before each treatment.

In order to guarantee anonymity, subjects were not asked to reveal their actual names or other information during the experiment, though we did have to collect and record names and payments for accounting purposes in the end. As mentioned, all monetary amounts in the experiment were denominated in ECUs with the rate of 100 ECUs = \$4.00. Each session lasted around 1 hour, with average earnings of \$25.6 per subject, including a \$10 show-up fee.

2.2.3 Hypotheses

Here are the hypotheses my experiment is designed to test in order to investigate employees' behavior in the new setting. Now employees in a group can be possibly affected by the feedback from the middle manager as before, and also by the feedback peer employees receive from the manager and efforts they exert if allowed to observe.

The following two hypotheses focus on treatment effects. In this experiment, as I illustrated, in addition to the Feedback treatment and the Feedback & Bonus treatment that might affect employees' effort choices, peer employees' feedback and efforts available in two treatment sessions can also influence employees' effort behavior. It is unknown whether employees are going to respond to these treatments, and it is also unclear how employees may shift their efforts in general. I state these treatments' corresponding hypotheses in positive narratives simply believing that they are able to help increase employees' efforts.

Hypothesis 2.1. Employees choose higher effort during the Peer Feedback session than under the Baseline session.

Hypothesis 2.2. Employees choose higher effort during the Peer Feedback & Effort session than under the Baseline session.

In my experiment, employees receive feedback that can be from 10 sad face emojis to 10 happy face emojis, and choose efforts from amount of 0 to 10. Besides the possible causal treatment effects brought by own feedback, peers' feedback, and peers' efforts, there can also exist correlations between employees' effort choices and the values of these variables[9]. For example, there can be a negative relationship between employees' own feedback and effort, which implies that employees can generally increase efforts after receiving bad feedback from middle managers, and reduce efforts when getting good feedback. Or a positive correlation between peers' efforts and employees' efforts can indicate that employees follow their peers' effort choices. These potential correlations may not be revealed by overall treatment effects, and I therefore present next three hypotheses to investigate employees' effort behavior in response to different values of own feedback, peers' feedback and peers' efforts.

Hypothesis 2.3. Employees choose higher effort in response to higher value of their own feedback from middle managers.

Hypothesis 2.4. Employees choose higher effort in response to higher value of their peers' feedback from middle managers.

Hypothesis 2.5. Employees choose higher effort in response to higher effort from their peer employees.

2.3 Results

I start to show the results of my experiment from the summary statistics. Table 2.2 presents a summary of employees' average wage, effort, feedback and frequency of bonus in different parts (treatments) and sessions. In the Baseline sessions, employees choose higher

---

[9]The value of feedback is defined as from -10 to 10 which corresponds to 10 sad face emojis to 10 happy face emojis.

average effort in part 3 of the experiment than pervious parts as they receive better feedback from middle managers and there is reward bonus allowed to be sent. In the Peer Feedback sessions and Peer Feedback & Effort sessions, there are no clear effort increases from employees on average in part 3 even though the reward bonus is still available, possibly due to the fact that they receive worse feedback from middle managers. Gift-exchange effects can be found in all three types of sessions since employees choose positive efforts rather than zero effort, but it does not imply that employees choose higher efforts when receiving higher wages, because employees' feedback and peers' feedback and efforts may have effects on their effort choices. For instance, compared to Peer Feedback sessions, employees receive higher wages in Peer Feedback & Effort sessions, but exert lower efforts.

	All	Part 1	Part 2	Part 3
(Baseline Session, 32 Subjects)				
Average Wage	14.28	14.28	14.28	14.28
Average Effort	4.44	4.13	4.17	5.03
Average Feedback	2.00	-	1.00	2.99
Frequency of Bonus	98	-	-	98
(Peer Feedback Session, 8 Subjects)				
Average Wage	15.00	15.00	15.00	15.00
Average Effort	5.48	5.30	5.9	5.25
Average Feedback	0.10	-	0.88	-0.68
Frequency of Bonus	20	-	-	20
(Peer Feedback & Effort Session, 8 Subjects)				
Average Wage	18.00	18.00	18.00	18.00
Average Effort	4.71	4.53	4.78	4.83
Average Feedback	-3.13	-	-2.78	-3.48
Frequency of Bonus	20	-	-	20

Table 2.2: Employees' Summary Statistics.

Figure 2.1 shows employees' average efforts by period in different sessions. Differences in average efforts can be observed in some parts, but again as there are varied wages and different kinds of treatments and sessions, a formal test to my hypotheses is still needed to figure out their effects on employees' effort choices.
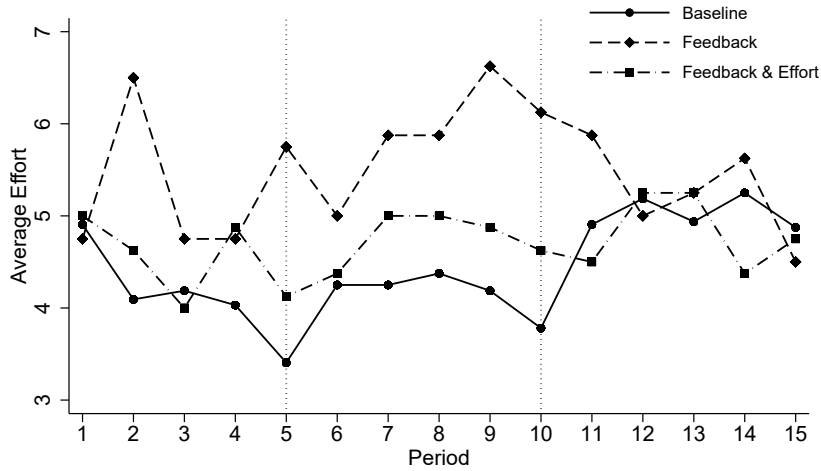
Figure 2.1: Average Effort in all periods, divided by Baseline, Peer Feedback and Peer Feedback & Effort sessions.

Table 2.3 includes a random effects regression of effort on wage, feedback, feedback & bonus, peer feedback and peer effort with standard errors clustered by subjects. In the regression, the feedback/feedback & bonus are treatment variables indicating whether employees are able to receive feedback/feedback & bonus or not from middle managers to themselves. Peer feedback is a treatment variable of whether employees can observe their peer employees' feedback or not, and peer effort is a treatment variable of whether employees can also observe their peer employees' efforts or not.

Based on the regression, there is still a significant positive correlation between employees' effort and wage which confirms the existence of gift-exchange effects. There is also a significant positive correlation between employees' effort and feedback & bonus. As for the rest treatments, there is no clear impact of them on employees' effort.

	Effort
Wage	0.339***
	(0.0707)
Feedback	0.171
	(0.173)
Feedback & Bonus	0.475***
	(0.167)
Peer Feedback	0.798
	(0.901)
Peer Effort	-1.793
	(1.283)
Constant	-0.678
	(0.971)
Observations	720
Number of SubjID	48

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.3: Random effects regression of effort on wage, feedback, feedback & bonus, peer feedback and peer effort.

Result 2.1. There is no significant difference in effort due to the Peer Feedback session.

Result 2.2. There is no significant difference in effort due to the Peer Feedback & Effort session.

As mentioned when raising hypotheses, even though there is no significant causal effect between employees' effort and feedback, peer feedback or peer effort as treatment, correlations between employees' efforts and the values of these feedback and effort can still exist. For example, it has been found in the study (Chen and Salmon, 2023) that employees might increase efforts in response to middle managers' negative feedback, while there lacks an overall treatment effect of middle managers' feedback on employees' efforts.

Figure 2.2 and Figure 2.3 also provide some hints about employees' effort changes when they know their peers' feedback and efforts. In Figure 2.2, when employees know that their own feedback from middle managers is lower than their peers' feedback, they tend to increase their efforts more likely, while when they realize their feedback is the same as or better than their peers' feedback, they seem to increase and decrease efforts equally. In Figure 2.3, as employees are able to check their peer employees' efforts, they increase their efforts more often if their efforts are lower than peers' efforts, and decrease their efforts more frequently

if their efforts are equal to or higher than peers' efforts.

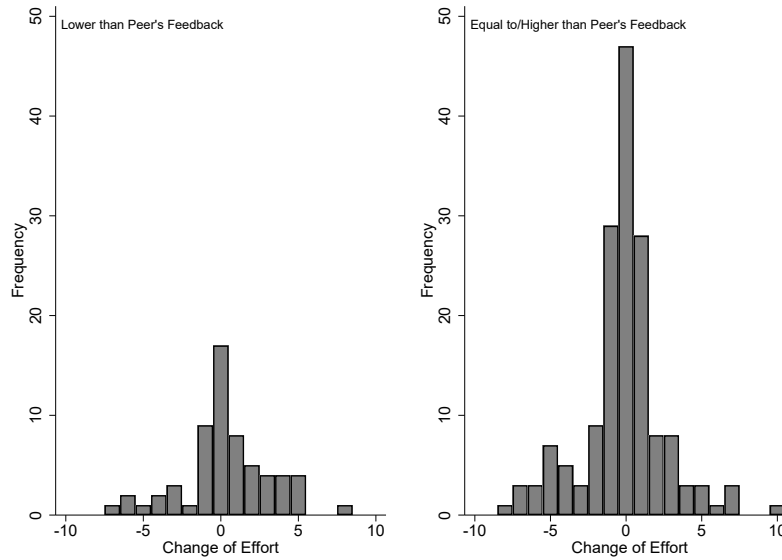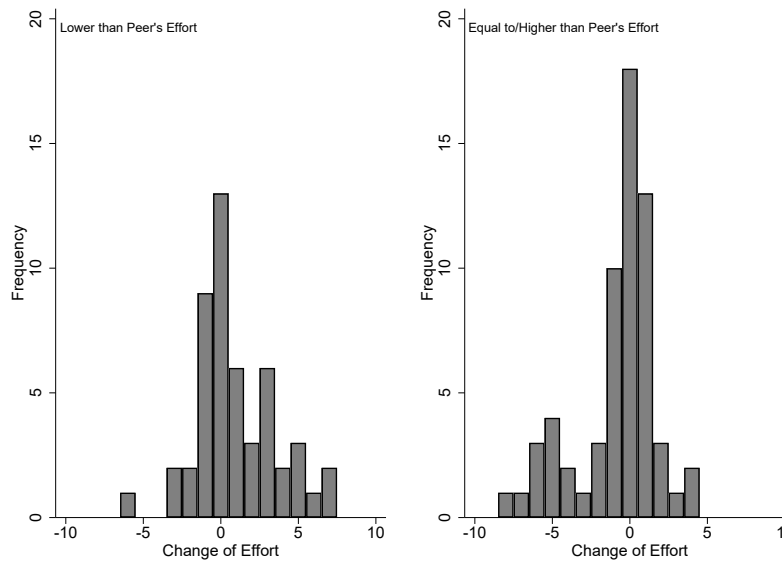Figure 2.2: Employees' changes of effort with comparison of feedback.

Figure 2.3: Employees' changes of effort with comparison of efforts.

Panel regressions are conducted to test the last three hypotheses. In Table 2.4, there are two random effects regressions with the same dependent variable change of effort, which is employees' change of efforts from period t-1 to t. The first regression includes independent

variables of wage, lagged value of feedback, lagged value of peer feedback and feedback & bonus. The value of feedback is the positive/negative number of the feedback employees receive from their middle managers in a period, and the value of peer feedback is the count of the feedback their peer employees receive in the same period. The second regression contains an additional independent variable lagged value of peer effort compared to the first one, which is the value of efforts from the peer employees in the same group. Standard errors are clustered by subjects. Both regressions show that there is a significant negative correlation between employees' change of effort and value of their own feedback, and as for the value of peer feedback, it has no significant impact on employees' change of effort. The second regression also indicates that there is significant positive relationship between employees' change of effort and value of peer effort.

	(1) Change of Effort	(2) Change of Effort
Wage	0.00882	-0.0143
	(0.0468)	(0.154)
Value of Feedback (t-1)	-0.148**	-0.250***
	(0.0702)	(0.0739)
Value of Peer Feedback (t-1)	0.0253	-0.149
	(0.0369)	(0.187)
Value of Peer Effort (t-1)		0.491*
		(0.267)
Feedback & Bonus	-0.440	-0.440
	(0.295)	(0.497)
Constant	-0.0829	-3.042
	(0.709)	(2.895)
Observations	144	72
Number of Subjects	16	8

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.4: (1) Random effects regressions of effort change on wage, lagged value of feedback, lagged value of peer feedback and feedback & bonus. (2) Random effects regressions of effort change on wage, lagged value of feedback, lagged value of peer feedback, lagged value of peer effort and feedback & bonus.

Result 2.3. There is a significant negative correlation between employees' efforts and their own feedback from middle managers.

Result 2.4. There is no significant difference of efforts if employees observe their peers' feedback from middle managers or not.

Result 2.5. There is a significant positive correlation between employees' efforts and their peer employees' efforts.

These results show that employees still present standard gift-exchange effects in the groups of multiple workers, as they choose to increase efforts in response to higher wages offered by middle managers. The Feedback & Bonus treatment is able to increase employees' efforts. Employees seem to react to the feedback of their own from middle managers but not the feedback of their peer workers, while they might raise efforts knowing coworkers' high efforts.

2.4   Conclusion

Peer effects may exist in an environment that includes multiple employees working with a middle manager and potentially change those employees' performance as they not only receive their own feedback as evaluations from the manager, but also be able to observe their peer workers' feedback from the manager and efforts. Previous study (Chen and Salmon, 2023) focusing on the relationship between a middle manager and one single employee provides persuasive results that employees can respond to middle managers' negative feedback by increasing their efforts, and the reward bonus with a small amount of monetary incentive is useful to uplift their efforts as well. But as there are now more than one middle manager-employee relationships, employees' behavior might alter because of these peer effects.

I implement an experiment to investigate how employees' effort behavior can be affected by those peer effects. The experiment is modified from the experiment with the basic middle manager-employee setting and now one extra employee is added in each group. It allows me to verify in the environment of one middle manager and two employees, if employees still respond to their own feedback from the manager in the way that the previous experiment suggests, and I am also able to find out whether and how employees react to the feedback their peer employees receive from the manager and efforts their peers choose if they can observe. This experiment can be considered as one of the first that include the information

of both performance feedback and efforts in a group of one middle manager and multiple employees and it therefore allows for an investigation of two types of peer effects in workplaces in a systematic way at the same time.

My current data from the experiment shows that, in the setting where employees can observe their peer workers' performance feedback and performance itself, the information of peer employees' performance feedback has no general impact on employees' effort choices, and the information of peer employees' performance also lacks an overall effect on employees' efforts. There is no treatment effect of employees' own performance feedback on their efforts as well. Some correlational effects, however, can be found in the data such as a negative relationship between employees' own performance feedback and efforts and a positive relationship between peer workers' performance and employees' effort choices, which suggests potential peer effects of performance. The lack of overall effect of own feedback on efforts can be due to the offset of increased and decreased efforts led by different levels of negative feedback. Similarly, employees might increase efforts when observing peer works' high efforts and decrease efforts when knowing coworkers' low performance and they can cancel out. As the result, there is also no overall treatment effect of peer employees' performance on employees' efforts.

The correlation between peers' efforts and employees' efforts echos the findings of some previous studies that positive peer effects of performance exist and drive workers to improve their efforts. As for the feedback from the middle managers, this study confirms the negative correlation between employees' efforts and own feedback, and reveals that peers' feedback has no impact on employees' effort decisions. Although only the peer effects of performance among employees are potentially found in my experiment, the peer effects of performance evaluations can still play some roles in influencing workers' effort choices simply in different means. For example, instead of just seeing peer workers' performance feedback, being able to communicate about their feedback from managers can possibly change employees' future performance even without the actual performance being disclosed during the communication. A public announcement of employees' performance evaluations by the middle manager

45

in a group may trigger more responses from employees despite the fact that they still only receive the information of peers' performance feedback. More methods allowing employees to exchange their performance feedback need to be tested before ruling out the effects of the feedback on employees' efforts.

In addition to various means of delivering peers' performance feedback among workers, a few more possible extensions can be considered as future studies. There exist a great deal of researches about the peer effects of workers' performance focusing on some other aspects. For instance, wage (Cornelissen et al., 2017; Hong and Lattanzio, 2022), different payment schemes (Eriksson et al., 2009; Bellemare et al., 2010; Milán and Oviedo-Dávila, 2024), perseverance and quitting (Buechel et al., 2018; Rosaz et al., 2016), team size and network (Chadi and Homolka, 2023; Lindquist et al., 2022), and observing/being observed (Georganas, 2015; Bönsch et al., 2017). Inspired by these studies, I can certainly continue my investigation of the peer effects of workers' performance feedback in different ways. In my experiment, the middle manager sets the same fixed wage to employees in a group and employees' efforts decide the managers' earnings independently. There can be some shifts of employees' effort behavior if their efforts together contribute to the middle manager's payoff, or the manager can set the wages to employees differently. I can also vary the payment schemes to the employees to observe their responses. The grouping is fixed in my experiment, but the employees' effort choices might change in the long term if they are allowed to quit. Team size and network may have some influences on the peer effects as well, since employees are able to observe more coworkers' performance feedback in a larger group, or they can check different coworker's information. Employees may behave differently if they are observing or being observed in workplaces, and I can therefore study the directions of the peer effects.

CHAPTER 3

Preferences for Avoiding Subjective Judgment

3.1 Introduction

Some certain groups of people in the U.S. are less represented in many occupations relative to their shares of the total population. Considering that women's share of the U.S. population is approximately 51%, and Hispanic, African American and Asian's shares are 19%, 13% and 6% respectively, women make up only 27% of all STEM workers in the U.S by 2019 (Martinez and Christnacht, 2021) and 23.5% of tenured and tenure-track professors or faculty members in economics (Bayer and Rouse, 2016). Only 6% and 5% of people in publishing-related works, such as editorials, sales and marketing, are Hispanic and black (Lee & Low Books, 2019). As for people working in legal services industry, 38.8% of them are women, 6.1% are Hispanic, 6.1% are black and 5.3% are Asian (U.S. Census Bureau, 2021).

There are a number of pre-existing explanations for the issue of underrepresentation. For example, it can be due to cultural beliefs or stereotypes about women's talents and skills (Steele, 1997; Reuben et al., 2014; Sheltzer and Smith, 2014), racial barriers that impact on experiences and equities in education (Quentin and Hermann, 2015; Beasley and Fischer, 2012; Ong, 2011), vertical and horizontal segregations for women among and in occupations (Anker, 1997; Maffía, 2008; Borrás and Bucci, 2016), implicit bias reflecting gender or race stereotypes (Bertrand and Mullainathan, 2004; Green et al., 2007), institutional discrimination with rules and practices (Ginther and Kahn, 2004; Lawless and Pearson, 2008) and fear of competition or negative feedback (Niederle and Vesterlund, 2007; Buser et al., 2014;

Chakraborty and Serra, 2021).

It is noticeable that industries such as finance and accounting have less significant underrepresentation problems than professions I have discussed: In 2019, 56.5% of financial managers are female, and Hispanic, African American and Asian's shares are 11.2%, 8.0% and 7.7%; women account for 59.6% of accountants and auditors in 2019, and Hispanic, African American and Asian have the shares of 8.4%, 9.1% and 12.2% (U.S. Census Bureau). One observation is that the work of finance and accounting can be evaluated in a more objective way, while in other professions mentioned above, people's works are more likely judged by some people's subjective criteria. For instance, an asset manager's compensation can be clearly illustrated and written down in a contract depending on the scale and return of the assets under her management. But for a literature author, although the reward is based on the sales of the book she writes, she first needs to pass the selection of editors in a publishing house, which is a process containing much higher degree of subjective evaluation due to the editors' own tastes and perspectives.

Most of the studies about underrepresentation and approaches to improve it, as mentioned, focus on the institutional demand-side, while as its consequence, the supply-side is less noticed. As these studies suggest that discrimination can take place in the workplaces, it is natural that females and minority people form their preferences for avoiding subjective judgment and therefore discrimination. In order to fix the problem of underrepresentation in some industries, people's avoidance to judgment needs to be paid attention to as well. I am therefore interested in for tasks whose rewards are determined by subjective judgment, if certain groups of people have reduced willingness to participate in just because there exists such a judgment? With the belief and even the experience of being mistreated and underrepresented, these people might expect that they are going to be judged with bias and then self-select to opt out from these tasks. For example, a minority individual may decide not to spend time into writing a book as she looks at the process of getting her book published through editors' judgment and selection, and instead work in a field where she doesn't have to go through the subjective screen.

I answer my research question by conducting a laboratory experiment. The experiment allows each participant to choose between a subjective task and an objective task, and then finish the task she chooses. In the subjective task, two participants answer the same family feud question. Their answers will be selected by a judge and the selected answer will be compared to answers to a separate survey. If a subject's answer belongs to the top three answers from the survey, the subject will get the reward. In the objective task, a participant answers a family feud question and the answer has a given probability to be delivered to the comparison. If the answer belongs to the top three answers from the survey, the subject will get the reward. By observing a participant's choices between pairs of subjective task and objective task with different probabilities from 20% to 70%, the participant's preference to the judgment can be identified when she switches from subjective tasks to objective tasks, as those family feud questions don't require any specific ability or skill, and the participant should expect 50% chance of being selected by the judge in subjective tasks if there is no bias in the selection. I am therefore able to find out if people have varied degrees of avoidance to judgment.

In the experiment, participants' identity information is designed to be available to the judge or not in two treatments, as I would like to investigate if people's avoidance to judgement can be affected by the salience of their background information. It would help alleviate their underrepresentation if females and minorities reduce their avoidance more than white people or males when their identity information is hidden to the judge, as I can then encourage them to participate in the tasks with subjective judgment by making their identities anonymous during the process.

My current data shows that there is no any significant correlation between participants' avoidance to subjective judgment and their genders, ethnicities, or demographic information being accessible to the judge, even though some indications can be observed that these preferences for avoiding subjective judgment might exist. Making subjects' identity information hidden cannot potentially help women and minorities' participation in subjective tasks, as

49

when the information is disclosed to the judge, no significant correlations between subjects' changes of avoidance to judgment and their genders or ethnicities are found.

3.2 Literature

A large number of studies try to present and explain certain groups of people's underrepresentation. There are a lot of papers showing that the bar women and minorities have to pass is higher than that of others. Ginther and Kahn (2004) find that when compared with other academic disciplines, women in economics are less likely to get tenure and take longer to achieve it. Lawless and Pearson (2008) analyze women's performance in general elections for the U.S. House of Representatives and conclude that due to the congressional primary process, women have to be "better" than their male counterparts in order to fare equally well. Factors such as financial issues and family considerations drive the schooling and careers choices of females and minorities. Feinberg (2020) finds that student debt, especially graduate student debt drives minorities out of university teaching and research positions. Ceci and Williams (2011) argue that women's underrepresentation is not because of the sex discrimination by grant agencies, journal reviewers, and search committees, while is due to factors surrounding family formation and childrearing, gendered expectations, lifestyle choices, and career preferences. The pipeline problem is also used to address the problem of underrepresentation, for example, women's absence of being in economics profession is partially because of the gender imbalance in undergraduate level of economics education (Emerson et al, 2012; Rask and Tiefenthaler, 2004; Dynan and Rouse, 1997). All these explanations can be concluded as the demand-side's implicit attitudes and institutional practices which are mentioned in the study of underrepresentation of women and minorities in the economics profession by Bayer and Rouse (2016), while the focus of my study is people's expectations of being treated with bias which belongs to the supply-side, as the result of the demand-side.

Niederle and Vesterlund (2007) find that women shy away from competition and men embrace it even though they have the same ability, for the reason that men are more overconfident and prefer to perform in a competition. Buser et al. (2014) further indicate that competitiveness is strongly positively correlated with choosing more prestigious academic

tracks even conditional on academic ability. Their studies, together with lots of studies showing gender differences of risk aversion (Eckel and Grossman, 2008), indicate that women and men do have various willingness to expose to tasks with competition or task. Here I try to find out if people with different demographics (not limited to gender) have divergent tendencies to enroll in tasks with subjective judgment.

Hoff and Pandey (2004) study how the past shapes belief systems that shape individuals' responses to opportunities. In their experiment, when caste identities are salient, low-caste individuals in India anticipate that they will be judged prejudicially and underperform. Ku and Salmon (2012) show that individuals exhibit a discouragement effect in the face of inequality that leads to lower work effort. Their studies focus on how people's performances are affected by the concern of being judged or discouraged, while my study is interested in people's choices and participations when anticipating judgment. My study relates to the context of discrimination, which has been studied by lots of scholars using experiments (Lane, 2016). This study also contributes to the literature of experimental studies of identity. As Charness and Chen (2020) summarize, one's perception of the social norm applicable to the salient identity decides behavior, which echos my study that subjects may self-select their tasks as they are aware of their identities and potential judgment.

3.3 Experiment

3.3.1 Design

At the beginning of the experiment, all the participants are asked to finish a basic demographic questionnaire including questions about each participant's age, gender and ethnicity.[10] There is no extra information provided about this questionnaire and participants will only know how their demographic information is used when it is the Information treatment.

In the experiment, there will be only one participant randomly assigned the role of judge, and all the rest of participants will be players. Their roles are notified once they finish the

---

[10]Each participants is able to choose an age between 18 and 80, a gender to be male, female and other, race and ethnicity

demographic questionnaire and fixed along the entire experiment. There are two parts (treatments) to the experiment, and each part contains three periods.

Each player in a period is going to make a choice between two tasks which are labeled as task 1 and task 2 for six different situations. So there are six choices to be made and one of them will be randomly realized to be finished by the player.

Prior to the experiment, an online survey was given to 400 people in Texas between the ages of 18 and 30, 200 men and 200 women. They were asked to provide their answer to many different family feud style questions such as "Name an instrument that is too big to carry on a plane". In both task 1 and task 2, the player will be asked to guess the most common answer to the questions the previous survey respondents gave. The difference between the two tasks is in how it is determined if the answer the player provide will be eligible for a reward.

If the player chooses task 1, she will be matched with another player to compete for a possible reward. Both of them will answer the same family feud style question. The participant in the role of the judge will be shown the answers submitted by both players in that pair. The judge will then select which of the two players will be eligible for a reward in that period. If the player is selected by the judge and her answer to the question is in the top two most common answers provided by the previous respondents, she will win a reward of $10. If the player is not selected or is selected but her answer is not among the top two answers, she will receive no reward in that period. For the individual in the role of the judge, she will be making selections for potentially multiple pairs of competitors in each period. After the judge has made selections for each pair, one of those pairs will be randomly picked to generate earnings for the judge. If the answer from the player the judge selected in that pair is in the top two answers, the judge will earn a reward of $5.

If the player chooses task 2, she will also be asked to provide her guess regarding the most common answer to a family feud style question. The player will not be matched against

another player and the human judge will not observe her answer. Instead there will be a clear and objective probability that will determine whether the player will be eligible for a reward. This probability will vary across the six situations she will be making choices for. In one case, the player would be designated as eligible for a reward with 20% probability, another with 30% and all the way up to 70%. If the player has chosen task 2, the computer will randomly determine whether she will be eligible for a reward according to the relevant probability. If the player is eligible and her answer is among the top two most common answers, she will receive a reward of $10.

In the Baseline treatment, players' demographic information is hidden to the judge, while in the Information treatment, their information is disclosed to the judge during her selections of players. Half of the experiment sessions will be implemented with the Baseline session first and then Information treatment, and the other half will be conducted in the reversed order, as I try to prevent any potential order effect.

Participants will not find out their rewards at the end of each period, and after all six periods a full accounting of results from each period and earnings will be provided. It is possible that there can be an odd number of players participating in task 1, and I would not be able to match all players in pairs. Consequently, one player will be chosen at random to switch to task 2 instead. In determining whether an answer matches with the top two from the survey, that check will be performed by the experimenter. This will allow us to still give credit for an answer in the event of a spelling mistake or a slightly different way of stating the same answer

3.3.2 Implementation

I conducted 7 experimental sessions at the Laboratory for Research in Experimental Economics (LREE) at Southern Methodist University and there were 62 participants in total. These 7 sessions included 4 Baseline-Information sessions with 30 subjects and 3 Information-Baseline sessions with 32 subjects. The experiment was programmed in z-Tree (Fischbacher, 2007) and subjects were recruited among pre-registered LREE students using an hroot based

recruitment system (Bock et al., 2014). Each subject participated in only one session and was grouped and assigned a role randomly.

I read aloud a script to explain laboratory rules and experiments general introduction after subjects were seated, and detailed instruction of a treatment before each part of the experiment. Subjects were also guided through sample decision-making screens before each treatment.

In order to guarantee anonymity, subjects were not asked to reveal their actual names or other information during the experiment, though I did have to collect and record names and payments for accounting purposes in the end. Each session lasted around 1 hour, with average earnings of $28.1 per subject, including a $10 show-up fee.

3.3.3 Hypotheses

My experiment is mainly designed to find out if minorities and females have reduced willingness to choose subjective tasks, and I have first two hypotheses referring to this central research question. In the experiment, a player chooses between task 1 competing with the other player and task 2 with a fixed probability. Since both tasks require players to answer a family feud question with no specific abilities needed, a player should be indifferent to task 1 and task 2 with 50% probability if the player doesn't have any bias toward the selection of the judge in task 1. However, players may choose task 1 less likely (be indifferent to task 2 with probability lower than 50%) if they have concerns about their identities and therefore the selections by the judge, and I can count the number of task 1 a player chooses to approximate the player's willingness to be judged.[11]

Hypothesis 3.1. Female subjects choose task 1 less often than male subjects.

Hypothesis 3.2. Minority subjects choose task 1 less often than white subjects.

---

[11]Here I count the number of task 1, instead of calculating the indifferent probability, as I allow the multi-switches between two tasks from players.

My next three hypotheses focus on the accessibility of the players' demographic information to the judge. When players' identity information is available to the judge during the selection, they may reduce their choices on task 1, and females and minorities can reduce even further if the concerns related to their identities exist. This can potentially suggest that hiding people's information is helpful to bring females and minorities to tasks with subjective judgment, if they are more sensitive to the salience of demographic information.

Hypothesis 3.3. When demographic information is disclosed, subjects decrease the choices of task 1.

Hypothesis 3.4. When demographic information is disclosed, female subjects decrease the choices of task 1 more than male subjects.

Hypothesis 3.5. When demographic information is disclosed, minority subjects decrease the choices of task 1 more than white subjects.

3.4 Results

Table 3.1 shows the summary statistics of players' average number of task 1, divided by gender, ethnicity and treatment. The table indicates that males present higher average number of task 1 than females and white people choose more task 1 than minority people on average. As for treatments, when their demographic information is available to view for the judge, all of the players reduce their choices of task 1. White and minority people's choices of task 1 decreases the same degree on average, while females reduce their average number of task 1 more than males.

	All	Male	Female	White	Minority
All Treatments	3.30	3.44	3.13	3.46	3.19
Baseline Treatment	3.38	3.49	3.22	3.54	3.26
Information Treatment	3.23	3.38	3.04	3.39	3.11

Table 3.1: Average number of players' choices on task 1.

55

Difference of choices of task 1 between males and females is shown in Figure 3.1. It can be observed that males choose more task 1 than females in two treatments, and they both reduce their choices of task 1 in Information treatment.
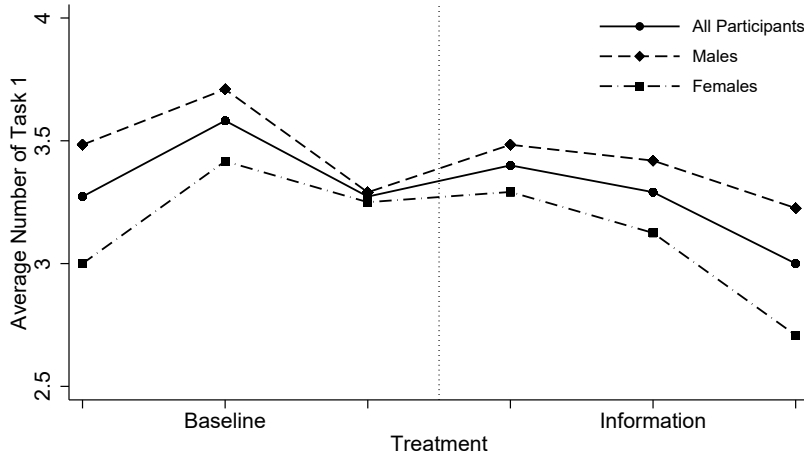

Figure 3.1: Average number of players' choices on task 1, divided by male and female subjects.

Figure 3.2 shows the numbers of choices of task 1 from white and minority people. As I can find in summary statistics, white people choose more task 1 than minorities, and both of them also avoid the task when their information is viewable to the judge.
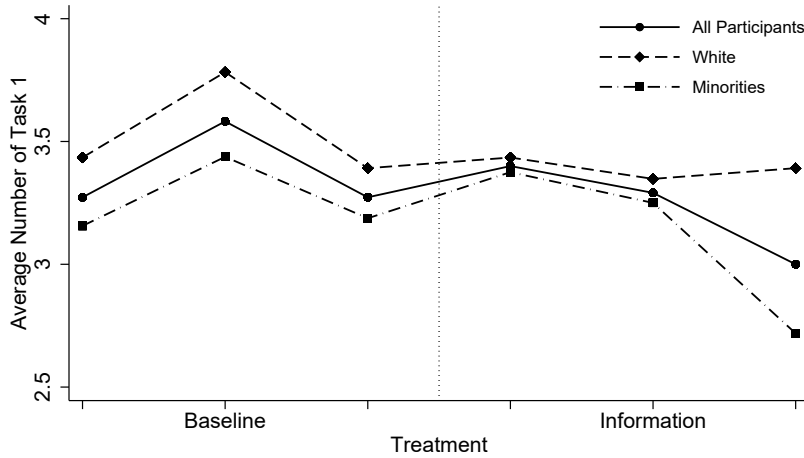
Figure 3.2: Average number of players' choices on task 1, divided by white and minority subjects.

Although the table and figures above suggest that different groups of people present varied behavior of choosing subjective task 1, I conduct a few regressions in order to formally test my hypotheses.

Table 3.2 includes a random effects regression in which the dependent variable is the number of task 1, and independent variables are age, female, minority, information, female $\times$ information and minority $\times$ information. Standard errors are clustered by subjects. Female, minority and information in the regression are treatment variables indicating if a player belongs to the group of females or not, the group of minorities or not, and if the player's information is revealed to the judge or not. Female/minority $\times$ information are treatment variables telling whether a female/minority player's information is disclosed to the judge or not.

Based on the regression, there is no significant correlation between players' choices of task 1 and gender, ethnicity or information, even though the regression shows the directions that females and minorities choose less task 1, and players choose less task 1 when the judge can access their information, which all follow the observation of the data. There is also no

significant correlation between female or minority players' choices of task 1 and their demographic information being salient to the judge.

	Number of Task 1
Age	-0.0572**
	(0.0289)
Female	-0.242
	(0.358)
Minority	-0.154
	(0.328)
Information	-0.123
	(0.257)
Female × Information	-0.0636
	(0.379)
Minority × Information	0.00877
	(0.379)
Constant	4.971***
	(0.853)
Observations	330
Number of Subjects	55

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.2: Random effects regression of number of task 1 on age, female, minority, information, female × information and minority × information.

Result 3.1. There is no significant correlation between players' choices of task 1 and their genders.

Result 3.2. There is no significant correlation between players' choices of task 1 and their ethnicities.

Result 3.3. There is no significant correlation between players' choices of task 1 and the disclosure of their demographic information.

I also run regressions to test last two hypotheses. Regressions in Table 3.3 are random effects regressions of number of task 1 on age, minority and information, divided by only males in the first regression and only females in the second regression. Standard errors are clustered by subjects. There is no significant correlation between the number of task 1 and information, no matter it is the male group or female group. As the result, when the infor-

mation is available to the judge, there is no significant correlation between players' changes of number of task 1 and their genders.

	(1) Number of Task 1 (Male)	(2) Number of Task 1 (Female)
Age	-0.0220	-0.0808*
	(0.0333)	(0.0413)
Minority	-0.163	-0.146
	(0.394)	(0.418)
Information	-0.118	-0.181
	(0.246)	(0.260)
Constant	4.117***	5.302***
	(0.987)	(1.066)
Observations	186	144
Number of Subjects	31	24

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Random effects regression of number of task 1 on age, minority and information, divided by male and female subjects.

Result 3.4. When demographic information is disclosed, there is no significant correlation between players' changes of choices of task 1 and their genders.

Table 3.4 shows two random effects regressions of number of task 1 on age, female and information, divided by only white subjects in the first regression and only minority subjects in the second regression. Standard errors are clustered by subjects. Still, I cannot find any significant correlation between the number of task 1 and information either in white group or minority group. Therefore I can't conclude that players' changes of number of task 1 correlate to their ethnicities, when their information is salient to the judge.

	(1)	(2)
	Number of Task 1 (White)	Number of Task 1 (Minority)
Age	-0.0658**	-0.0472
	(0.0322)	(0.0494)
Female	-0.236	-0.304
	(0.436)	(0.425)
Information	-0.145	-0.146
	(0.262)	(0.243)
Constant	5.171***	4.597***
	(1.003)	(1.314)
Observations	138	192
Number of Subjects	23	32

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Random effects regression of number of task 1 on age, female and information, divided by white and minority subjects.

Result 3.5. When demographic information is disclosed, there is no significant correlation between players' changes of choices of task 1 and their ethnicities.

It is possible that the reason of these correlations not being significant is that I don't have a larger number of subjects in my experiment. I therefore regroup my subjects to be minority and female or not as an indication for whether or not the lack of significance is simply due to the small sample size.

Table 3.5 shows a random effects regression of number of task 1 on age, minority & female, information and minority & female × information with standard errors clustered by subjects. Minority & female is a treatment variable indicating if a player falls into the category of minority and female or not. Minority & female × information is a treatment variable telling whether a minority female player's information is disclosed to the judge or not. I can't find a significant correlation between players' choices of task 1 and them belonging to the group of minority and female.

	Number of Task 1
Age	-0.0570**
	(0.0288)
Minority & Female	-0.561
	(0.405)
Information	-0.282
	(0.223)
Minority & Female × Information	0.470
	(0.337)
Constant	4.933***
	(0.783)
Observations	330
Number of Subjects	55

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Random effects regression of number of task 1 on age, minority & female, information and minority & female × information.

Table 3.6 has two random effects regressions of number of task 1 on age and information, divided by only minority and female in the first regression and other in the second regression. Standard errors are clustered by subjects. Just as I have seen in Table 3.3 and Table 3.4, players don't significantly change their numbers of task 1 as their information is viewable to the judge regardless they are in the minority & female group or not. So when demographic information is disclosed to the judge, players' changes of number of task 1 are not correlated to them being minorities and females.

	(1)	(2)
	Number of Task 1 (Minority & Female)	Number of Task 1 (Other)
Age	-0.0469	-0.0621**
	(0.0617)	(0.0306)
Information	0.187	-0.282
	(0.259)	(0.224)
Constant	4.119***	5.057***
	(1.360)	(0.854)
Observations	96	234
Number of Subjects	16	39

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Random effects regression of number of task 1 on age and information, divided by minority & female and other subjects.

I am also able to calculate the number of subjects needed so that the correlation in hypothesis 3.1 above can be significant, using a two-sample means test in a cluster randomized

design. Given average numbers of task 1 from male and female groups being 3.44 and 3.13, with common standard deviation of 1.64 and cluster sizes of 6 using default intraclass correlation of 0.5, power of 0.8, and significance level $\alpha = 0.05$, the number of subjects I need to generate a significant finding is 514. Similarly, to make the correlations in hypothesis 3.2 and hypothesis 3.3 to be significant, I need around 676 and 2,190. To be able to make the correlations in hypothesis 3.4 and hypothesis 3.5 to be significant, I need 8,466 and 304,764 subjects. Sample sizes of this magnitude are well beyond what would be feasible for a lab study such as this experiment and suggest that the null finding is a true null result rather than simply a matter of a small sample size.

3.5   Conclusion

This study originates from an observation that underrepresentation of females and ethnic/racial minorities exists in different occupations with varied degrees, which depends on the nature of the evaluations of these jobs. Professions with tasks including higher degree of objective criteria are more favored by females and minorities, and they present decreased willingness to joining in those with tasks requiring more subjective judgment. It can be due to their preferences for avoiding subjective judgment that drives these people to choose to participate in occupations differently. A great deal of previous studies try explain and improve the underrepresentation problem on the institutional demand-side, but people's preferences of avoidance to subjective judgment on the supply-side, as the result of existing bias and discrimination in workplaces, are always ignored. It is necessary to verify the existence of such preferences, for the reason that they may also contribute to the underrepresentation problem which therefore cannot be fixed simply by removing those bias and discrimination. Taking people's potential preferences for avoiding subjective judgment into account is an important contribution of my study to the literature about discrimination and underrepresentation in workplaces.

I implement an experiment to examine people's preferences of avoidance to subjective judgment. The experiment characterizes a scenario that participants are able to choose between tasks that only differ in the process of evaluation and selection: one with a subjective

judgment and the other with an objective probability. It allows me to test if people belonging to groups of females or minorities show higher preferences to avoid subjective judgment. I also include treatments that participants' demographic information can be viewed or not by the person who judges their tasks, so I am able to find out if females and minorities avoid subjective judgment more than males and white people when their information is disclosed, and therefore suggest that making their information hidden during the process of evaluation and selection can help increase females and minorities' participation in certain occupations.

The data from my experiment shows a lack of any significant correlation between participants' avoidance to subjective judgment and their genders, ethnicities, or demographic information being salient to the judge. There is also no significant correlation between subjects' changes of avoidance to judgment and their genders or ethnicities, when demographic information can be viewed by the judge.

Despite the fact that the correlations above are found to be insignificant in my experiment, there do exist some indications from the data like females and minorities present reduced willingness to participating in tasks with subjective judgment, and people increase their avoidance to subjective tasks when their demographic information is salient to the judge. This suggests that such preferences may exist but they are not triggered by the stylized nature of this specific experiment. This experiment was designed as a very strong test of these preferences as there are not many contextual elements which might have triggered such preferences and these results may simply suggest that absent such triggers, these groups do not avoid subjective judgment. This leads to a potential future research agenda to determine what other contextual elements might be necessary to add in to observe stronger effects as a way of determining if these preferences should or should not be expected to influence choices in the field.

First, my experiment displays six pairs of tasks for players to make choices in a screen adopting the method Holt and Laury (2002) use for eliciting risk preferences. Players select one task which can be either task 1 or task 2 with a given probability in each pair. They are

63

expected to choose task 1s in the first several pairs as corresponding task 2s have low probabilities and then change to task 2s in the rest pairs with task 2s having high probabilities. However, from the data, I find that a few players choose task 2s first and then task 1s, or present multi-switches across two tasks. For a relatively small sample size, multi-switches need to be ruled out so in the screen players will only report the pair in which they are going to change their task choices. Read script and screen design can be revised to help players better understand the choices and interactions they are going to make in the experiment.

Second, players in the experiment are asked to answer family feud style questions in both task 1 and task 2. These questions are designed to ask players' general knowledge of common things so there are no specific abilities needed from them and they therefore have no bias towards the questions and their identities. As the result, in my experiment, female and minority players can be more confident about their performance and the judge's selections so they present less avoidance to the judgment. However, in the real world cases, the judgment can be requiring certain abilities and time-consuming, and they may actually avoid the judgment more likely. This also applies to the identity of the judge. In the field cases, people can expect and meet a judge to be a white male more often, while in my experiment, the judge can be anyone in the lab whose identity is unknown to the players. They might only focus on the questions rather than the process of the judge making judgment and selections. In my future study, tasks and the process of judging and selecting in the experiment can include some aspects of discrimination that indeed exist outside the lab and people have awareness of.

Third, the lab experiment itself may reduce participants' expectation of being treated differently. As an academic experiment, it is formally implemented from subjects' recruitment to participation. Subjects are told and understand that it is a scientific study in a lab on campus and their privacies are respected and protected. What they have gone through all along the experiment leaves them an impression that the experiment treats them equally as participants, although the content of which relates to some degrees of discrimination on them. Subjects may experience discrimination outside of the lab, but may not be primed to expect differential treatment in such an academic experiment. A future experiment could

find a more realistic environment prone to discrimination where certain groups of people might normally expect disadvantageous treatment.

REFERENCES

[1] D. Sull, C. Sull, and B. Zweig, "Toxic Culture Is Driving the Great Resignation," *MIT Sloan Management Review*, 2022.

[2] S. Fox and L. E. Stallworth, "Racial/ethnic Bullying: Exploring Links between Bullying and Racism in the US Workplace," *Journal of Vocational Behavior*, vol. 66, pp. 438–456, 2005.

[3] M. B. Nielsen, S. B. Matthiesen, and S. Einarsen, "The Impact of Methodological Moderators on Prevalence Rates of Workplace Bullying. A Meta-analysis," *Journal of Occupational and Organizational Psychology*, vol. 83, pp. 955–979, 2010.

[4] H. Hoel, K. Sparks, and C. L. Cooper, "The Cost of Violence/stress at Work and the Benefits of a Violence/stress-free Working Environment," *Geneva: International Labour Organization*, 2001.

[5] P. Lutgen-Sandvik, S. J. Tracy, and J. K. Alberts, "Burned by Bullying in the American Workplace: Prevalence, Perception, Degree and Impact," *Journal of Management Studies*, vol. 44, no. 6, pp. 837–862, 2007.

[6] T. P. Index, "2022 Report: The State of Talent Optimization," 2022.

[7] T. S. for Human Resource Management, "The High Cost of a Toxic Workplace Culture," 2019.

[8] E. Dodds, "The 'Death Penalty' and How the College Sports Conversation Has Changed," *Time*, 2015.

[9] L. Wertheim, "All the Rage: Bobby Knight's infamous 'Chair Game', 30 years later," *Sports Illustrated*, 2015.

[10] M. Krzyzewski, "Coach K on How to Connect," *Wall Street Journal*, 2001.

[11] N. Mertens, F. Boen, G. V. Broek, M. Vansteenkiste, and K. Fransen, "An Experiment on the Impact of Coaches' and Athlete Leaders' Competence Support on Athletes' Motivation and Performance," *Scandinavian Journal of Medicine & Science in Sports*, vol. 28, no. 12, pp. 2734–2750, 2018.

[12] D. Farnsworth, J. L. Clark, A. Wysocki, and K. Kepner, "Managerial Feedback, Associate Performance, and Eleven Positive Feedback Rules," *HR026, the Food and Resource Economics Department, UF/IFAS Extension*, 2020.

[13] A. S. Hirsch, "Mastering the Art of Negative Feedback," *The Society for Human Resource Management*, 2018.

[14] R. Peeters and M. Vorsatz, "Immaterial Rewards and Sanctions in a Voluntary Contribution Experiment," *Economic Inquiry*, vol. 51, no. 2, pp. 1442–1456, 2013.

[15] S. Dugar, "Non-monetary Incentives and Opportunistic Behavior: Evidence from A Laboratory Public Good Game," *Economic Inquiry*, vol. 51, no. 2, pp. 1374–1388, 2013.

[16] S. Kandul and B. Lanz, "Public Good Provision, In-group Cooperation and Out-group Descriptive Norms: A Lab Experiment," *Journal of Economic Psychology*, vol. 85, no. C, 2021.

[17] E. Fehr, G. Kirchsteiger, and A. Riedl, "Does Fairness Prevent Market Clearing? An Experimental Investigation," *The Quarterly Journal of Economics*, vol. 108, no. 2, pp. 437–459, 1993.

[18] D. Masclet, C. Noussair, S. Tucker, and M.-C. Villeval, "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *The American Economic Review*, vol. 93, no. 1, pp. 366–380, 2003.

[19] M. Lefebvre and A. Stenger, "Short- & Long-term Effects of Monetary and Nonmonetary Incentives to Cooperate in Public Good Games: An Experiment," *Public Library of Science One*, vol. 15, no. 1, pp. 1–17, 2020.

[20] D. J. Cooper and J. P. Lightle, "The Effect of Structured Emotion Expression on Reciprocity in Bilateral Gift Exchange," *Experiments in Organizational Economics*, pp. 1–18, 2016.

[21] U. Gneezy and A. Rustichini, "Pay Enough or Don't Pay at All," *The Quarterly Journal of Economics*, vol. 115, no. 3, pp. 791–810, 2000.

[22] R. M. Titmuss, "The Gift Relationship: From Human Blood to Social Policy," *Allen and Unwin*, 1970.

[23] C. Mellstrom and M. Johannesson, "Crowding out in Blood Donation: Was Titmuss Right?," *Journal of the European Economic Association*, vol. 6, pp. 845–863, 2008.

[24] U. Fischbacher, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, vol. 10, pp. 171–178, 2007.

[25] O. Bock, I. Baetge, and A. Nicklisch, "hroot: Hamburg Registration and Organization Online Tool," *European Economic Review*, vol. 71, no. C, pp. 117–120, 2014.

[26] H. Chen and T. Salmon, "The Plight of the Middle Manager," 2023.

[27] C. Welteke, "Peers at Work - a Brief Overview of the Literature on Peer Effects at the Workplace and the Policy Implications," *DIW Roundup: Politik im Fokus*, vol. 68, 2015.

[28] M. C. Villeval, "Performance Feedback and Peer Effects: A Review," 2020.

[29] M. D. Paola, R. Nisticò, and V. Scoppa, "Workplace Peer Effects in Fertility Decisions," *Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy*, vol. 714, 2024.

[30] I. Ayres, S. Raseman, and A. Shih, "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage," *The Journal of Law, Economics, and Organization*, vol. 29, no. 5, pp. 992–1022, 2013.

[31] A. Mas and E. Moretti, "Peers at Work," *American Economic Review*, vol. 99, no. 1, pp. 112–145, 2009.

[32] A. Falk and A. Ichino, "Clean Evidence on Peer Effects," *Journal of Labor Economics*, vol. 24, no. 1, 2006.

[33] J. Guryan, K. Kroft, and M. J. Notowidigdo, "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, vol. 1, no. 4, pp. 34–68, 2009.

[34] C. S. Fan, X. Wei, J. Wu, and J. Zhang, "Observability and Peer Effects: Theory and Evidence from a Field Experiment," *Journal of Economic Behavior & Organization*, vol. 200, pp. 847–867, 2022.

[35] R. van Veldhuizen, H. Oosterbeek, and J. Sonnemans, "Peers at Work: Evidence from the Lab," *PLoS One*, vol. 13, no. 2, 2018.

[36] D. J. Zimmerman, "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment," *The Review of Economics and Statistics*, vol. 85, no. 1, pp. 9–23, 2003.

[37] E. Duflo, P. Dupas, and M. Kremer, "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, vol. 101, no. 5, pp. 1739–1774, 2011.

[38] J. J. Horton, "Employer Expectations, Peer Effects and Productivity: Evidence from a Series of Field Experiments," *arXiv preprint arXiv:1008.2437*, 2010.

[39] T. Cornelissen, C. Dustmann, and U. Schönberg, "Peer Effects in the Workplace," *American Economic Review*, vol. 107, no. 2, pp. 425–456, 2017.

[40] L. Hong and S. Lattanzio, "The Peer Effect on Future Wages in the Workplace," 2022.

[41] T. Eriksson, A. Poulsen, and M. C. Villeval, "Feedback and Incentives: Experimental Evidence," *Labour Economics*, vol. 166, no. 6, pp. 679–688, 2009.

[42] C. Bellemare, P. Lepage, and B. Shearer, "Peer Pressure, Incentives, and Gender: An Experimental Analysis of Motivation in the Workplace," *Labour Economics*, vol. 7, no. 1, pp. 276–283, 2010.

[43] P. Milán and N. O. Dávila, "Incentive Contracts and Peer Effects in the Workplace," *Barcelona School of Economics*, vol. 1439, 2024.

[44] B. Buechel, L. Mechtenberg, and J. Petersen, "If I Can Do It, So Can You! Peer Effects on Perseverance," *Journal of Economic Behavior & Organization*, vol. 155, pp. 301–314, 2018.

[45] J. Rosaz, R. Slonim, and M. C. Villeval, "Quitting and Peer Effects at Work," *Labour Economics*, vol. 39, pp. 55–67, 2016.

[46] A. Chadi and K. Homolka, "Under (peer) pressure: Experimental Evidence on Team Size and Task Performance," *Managerial and Decision Economics*, vol. 44, no. 7, pp. 3769–3786, 2023.

[47] M. J. Lindquist, J. Sauermann, and Y. Zenou, "Peer Effects in the Workplace: A Network Approach," 2022.

[48] S. Georganas, M. Tonin, and M. Vlassopoulos, "Peer Pressure and Productivity: The Role of Observing and Being Observed," *Journal of Economic Behavior & Organization*, vol. 117, pp. 223–232, 2015.

[49] A. Bönsch, J. Wendt, H. Overath, Ö. Gürerk, C. Harbring, C. Grund, T. Kittsteiner, and T. W. Kuhlen, "Peers at Work: Economic Real-effort Experiments in the Presence of Virtual Co-workers," *IEEE Virtual Reality (VR)*, pp. 301–302, 2017.

[50] A. Martinez and C. Christnacht, "Women Are Nearly Half of U.S. Workforce but Only 27% of STEM Workers," 2021.

[51] A. Bayer and C. E. Rouse, "Diversity in the Economics Profession: A New Attack on an Old Problem," *Journal of Economic Perspectives*, vol. 30, no. 4, pp. 221–242, 2016.

[52] Lee and L. Books, "Diversity in Publishing," *Diversity Baseline Survey 2.0*, 2019.

[53] U. C. Bureau, "ACS PUMS 1-Year Estimate," 2021.

[54] C. M. Steele, "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance," *American Psychologist*, vol. 52, no. 6, pp. 613–629, 1997.

[55] E. Reuben, P. Sapienza, and L. Zingales, "How Stereotypes Impair Womens Careers in Science," *Proceedings of the National Academy of Sciences*, vol. 111, no. 12, pp. 4403–4408, 2014.

[56] J. M. Sheltzer and J. C. Smith, "Elite Male Faculty in the Life Sciences Employ Fewer Women," *Proceedings of the National Academy of Sciences*, vol. 111, no. 28, pp. 10107–10112, 2014.

[57] Q. R. Alexander and M. A. Hermann, "African-American Womens Experiences in Graduate Science, Technology, Engineering, and Mathematics Education at a Predominantly White University: A Qualitative Investigation," *Journal of Diversity in Higher Education*, vol. 9, no. 4, pp. 307–322, 2016.

[58] M. A. Beasley and M. J. Fischer, "Why They Leave: The Impact of Stereotype Threat on the Attrition of Women and Minorities from Science, Math and Engineering Majors," *Social Psychology of Education: An International Journal*, vol. 15, no. 4, pp. 427–448, 2012.

[59] M. Ong, C. Wright, L. Espinosa, and G. Orfield, "Inside the Double Bind: A Synthesis of Empirical Research on Undergraduate and Graduate Women of Color in Science, Technology, Engineering, and Mathematics," *Harvard Educational Review*, vol. 81, no. 2, pp. 172–209, 2011.

[60] R. Anker, "Theories of Occupational Segregation by Sex: An Overview," *International Labour Review*, vol. 136, pp. 315–339, 1997.

[61] D. Maffía, "Carreras de obstculos, las mujeres en ciencia y tecnologa," *Paper presented at the Seminar on Women Scientists in Havana*, 2008.

[62] G. Borrás and I. Bucci, "Estudio, trabajo y discriminacin de gnero," 2016.

[63] M. Bertrand and S. Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, vol. 94, no. 4, pp. 991–1013, 2004.

[64] A. R. Green, D. R. Carney, D. J. Pallin, L. H. Ngo, K. L. Raymond, L. I. Iezzoni, and M. R. Banaji, "Implicit Bias among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients," *Journal of General Internal Medicine*, vol. 22, no. 9, p. 12311238, 2007.

[65] D. K. Ginther and S. Kahn, "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?," *Journal of Economic Perspectives*, vol. 18, no. 3, pp. 193–214, 2004.

[66] J. L. Lawless and K. Pearson, "The Primary Reason for Women's Underrepresentation? Reevaluating the Conventional Wisdom," *The Journal of Politics*, vol. 7, no. 1, pp. 67–82, 2008.

[67] M. Niederle and L. Vesterlund, "Do Women Shy Away From Competition? Do Men Compete Too Much?," *The Quarterly Journal of Economics*, vol. 122, no. 3, pp. 1067–1101, 2007.

[68] T. Buser, M. Niederle, and H. Oosterbeek, "Gender, Competitiveness, and Career Choices," *The Quarterly Journal of Economics*, vol. 129, no. 3, pp. 1409–1447, 2014.

[69] P. Chakraborty and D. Serra, "Gender and Leadership in Organizations: Promotions, Demotions and Angry Workers," *Working Papers 20210104-001, Texas A&M University, Department of Economics*, 2021.

[70] R. M. Feinberg, "Is an Academic Career a Luxury Good? Student Debt and the Under- representation of Minorities," *Economics Bulletin, AccessEcon*, vol. 40, no. 4, pp. 2964–2977, 2020.

[71] S. J. Ceci and W. M. Williams, "Understanding Current Causes of Women's Underrepresentation in Science," *Proceedings of the National Academy of Sciences*, vol. 108, no. 8, pp. 3157–3162, 2011.

[72] T. L. N. Emerson, K. McGoldrick, and K. J. Mumford, "Women and the Choice to Study Economics," *The Journal of Economic Education*, vol. 43, no. 4, pp. 349–362, 2012.

[73] K. Rask and J. Tiefenthaler, "Too Few Women? - or Too Many Men? The Gender Imbalance in Undergraduate Economics," *SSRN Electronic Journal*, 2004.

[74] K. E. Dynan and C. E. Rouse, "The Underrepresentation of Women in Economics: A Study of Undergraduate Economics Students," *The Journal of Economic Education*, vol. 28, no. 4, pp. 350–368, 1997.

[75] C. Eckel and P. Grossman, "Men, Women and Risk Aversion: Experimental Evidence," *Handbook of Experimental Economics Results*, vol. 1, no. 113, pp. 1061–1073, 2008.

[76] K. Hoff and P. Pandey, "Belief Systems and Durable Inequalities: An Experimental Investigation of Indian Caste," 2004.

[77] H. Ku and T. C. Salmon, "The Incentive Effects of Inequality: An Experimental Investigation," *Southern Economic Journal*, vol. 79, no. 1, pp. 46–70, 2012.

[78] T. Lane, "Discrimination in the Laboratory: A Meta-analysis of Economics Experiments," *European Economic Review*, vol. 90, pp. 375–402, 2016.

[79] G. Charness and Y. Chen, "Social Identity, Group Behavior, and Teams," *Annual Review of Economics*, vol. 12, pp. 691–713, 2020.

[80] C. A. Holt and S. K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, vol. 92, no. 5, pp. 1644–1655, 2002.