Journal of Behavioral and Experimental Economics 74 (2018) 1–16

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

Relative performance feedback: Effective or dismaying?

William Gilje Gjedrem

UiS Business School, University of Stavanger, 4036 Stavanger, Norway

A R T I C L E I N F O

Keywords:
Feedback
Personnel
Motivation
Productivity
Experiment
Incentives

JEL classification:
J24
M52
J33
C91
D23

A B S T R A C T

In this experiment, I analyze whether the provision of relative performance feedback differently affects the performance of subjects when provided in various feedback environments. Subjects were ranked either relative to the performance of many subjects in the past or relative to three subjects working alongside themselves. Results indicate that the response from subjects in the former varies with how they perceived their own ability to solve the task. Those reporting low ability reduce their performance when provided with the feedback, whereas those reporting high ability improve. For subjects who were ranked relative to others working alongside themselves, no one respond negatively, but only those reporting high ability improve their performance. An important implication from this, especially for managers who design feedback policies in organizations, is that the way relative feedback is designed may lead to different behavioral reactions. In particular, the choice of benchmark used to relatively rank employees may result in responses that are not beneficial and lead to inefficient use of resources.

1. Introduction

Information technology has made it easier for firms to evaluate employee performance more precisely and to use these evaluations to rank employees in relation to each other. It might be tempting for firms to (uncritically) adopt these modern evaluation tools, believing that it will boost performance to new heights. Understanding the full extent of how relative performance feedback (hereafter feedback) affects employees is complex, as competing social mechanisms are likely to influence employees simultaneously. For example, while competition between employees may lead them to exert higher effort, it may also make them feel incompetent. A particular worry is that some mechanisms “crowd out” employees’ intrinsic motivation to work (Deci, 1971; Frey and Oberholzer-Gee, 1997).

Two aspects of peoples’ social concerns are likely to be important reasons why feedback affect motivation; they have competitive preferences and people care about whether they feel competent or not. The latter aspect is considered the core of intrinsic motivation (Deci and Ryan, 2000), and learning about the performance of others may adjust the perception of own competence. However, people have competitive preferences too, which strengthens with the introduction of relative performance feedback. These competitive preferences may arise from reasons such as joy of outperforming others (Dohmen et al., 2011) or a desire for public recognition. In an effort to disentangle the effects of feedback these social concerns, this paper presents an experiment with treatments designed to feature conditions for competition and ability comparison separately.

In a lab experiment, two treatments are designed to feature each social concern separately. The first treatment, referred to as the CPF treatment, uses others’ past performance as benchmark to rank the current subject’s performance. Importantly, subjects in this treatment do not learn anything about the performance of any other subject in the same session. Thus, the environment is designed to reduce the competitiveness to a minimum, and should rather provide a signal about the general competence level of others to solve the task. The second treatment, referred to as the TPF treatment, uses the performance of three others working alongside the subject as benchmark for ranking. This should raise the competitiveness to a higher level, as subjects directly compete against each other for high ranks. In contrast to the former treatment, there are only noisy signals about the general competence level of others. These two treatments are compared to a baseline in which subjects only learn information about their individual performance (absolute performance feedback). In addition to these three different feedback conditions, treatments are also varied across fixed pay and performance pay. Subjects work on a real effort task. Before being provided with any feedback, they are asked to self-assess their own ability to solve such tasks. The perception of own ability may prove important to future responses to feedback (Gibbons and McCoy, 1991; Abeler et al., 2011).

The overall results, using non-parametric tests, suggest no

E-mail address: william.g.gjedrem@uis.no.

https://doi.org/10.1016/j.socec.2018.02.008
Received 15 July 2016; Received in revised form 20 February 2018; Accepted 21 February 2018
Available online 22 February 2018
2214-8043/

performance difference between the baseline and treatments under any pay scheme. However, regression analysis is required to adequately control for subjects' ability and to test for heterogeneous reactions. These analyses show that, when payment is fixed, the average performance of subjects is greater in both treatments compared to the baseline, but this is only significant in the CPF treatment. Large variations in performance exist, especially in the CPF treatment where subjects who report low ability reduce their performance substantially when feedback is provided. For the equivalent group of subjects in the TPF treatment, no such negative response has been identified. Moreover, those with high self-assessed ability (SAA) perform better in both treatments. However, further analysis show that those who perform worse than their expected ability level actually are the subjects who improve their performance the most in the CPF treatment. In the performance pay conditions of the experiment, no average treatment effects have been identified.

This study has two main contributions. First, it analyzes the effects of feedback in environments tailored to separately feature either comparison of ability or competition, which are considered to be important reasons to why feedback affects people. This makes it possible to isolate the effects of feedback when ability comparison is facilitated and isolate the effects of feedback when competition is facilitated, providing more insight to why people respond as they do to feedback. This should provide insightful information to managers in organizations, as they have numerous ways (frequency, benchmark, etc.) to design feedback in their organization. Their choices affect the degree of competition that the feedback will generate, and to which extent it enables colleagues to assess their competence. Both factors are likely to affect the net outcome on performance. Second, it shows that the belief about own ability may play a key role to how people respond to feedback, and that there is a worry of effort distortion for those who are in the lower tail of the performance distribution when providing feedback. Whether you consider yourself competent or not at a particular task, may influence how susceptible you are to learn about the performance of others.

This paper relates to a series of empirical studies on feedback in recent years, both from the lab and in the field. Lab settings that are closely related to this paper includes Hannan et al. (2008), Eriksson et al. (2009), Charness et al. (2014), and Azmat and Iriberri (2016). Using a real effort task, Eriksson et al. (2009) did not find any change in performance as a result of providing feedback under piece rate, although subjects made more mistakes. Charness et al. (2014) find higher performance amongst subjects who received feedback under fixed pay. In Azmat and Iriberri (2016), subjects only improved performance when paid a piece rate, whereas performance remained unchanged under fixed rate. Similarly, Hannan et al. (2008) find that performance improves under piece rate, is unchanged under fixed rate, and is reduced in a tournament setting. A series of other lab papers (Murthy and Schafer, 2011; Kuhnen and Tymula, 2012; Tafkov, 2012; Hannan et al., 2013) have primarily shown positive performance effects of providing feedback in various settings. Most relevant for this paper is Kuhnen and Tymula (2012) who find that the impact of feedback depends on the subjects' prior beliefs, and that subjects who rank lower than expected increase effort and vice versa. Moreover, some have considered tournaments between agents, in which some subjects also play the principal role of providing feedback (Ederer and Fehr, 2007; Gürtler and Harbring, 2010). In this setting, the principal sometimes provide untruthful information to the agents. Finally, in a recent paper by Gjedrem and Kvaløy (2016), there are no performance difference when subjects receive feedback and are working under piece rate, and the same applied to subjects working in teams and receiving relative feedback on team level.

In field settings, relevant papers have been conducted both in school and in the workplace. Bandiera et al. (2009) and Azmat and Iriberri (2010) find that students improve their subsequent performance after learning about the performance of other students, whereas Azmat et al. (2016) find a short lasting decrease in performance when feedback is provided to students. In workplace settings, both Blanes i Vidal and Nossol (2011) and Barankay (2012) relates to this paper. In Blanes i Vidal and Nossol (2011), employees improve productivity by almost 7% when they started to receive feedback, and this effect did not fade out during the period of study. Barankay (2012), on the other hand, finds that sales staff in a furniture retailer increase performance after feedback is removed, even though pay was not linked to relative performance. Finally, feedback at team level in the workplace have shown that ranking incentives may reduce the performance of the lowest ranked teams (Bandiera et al., 2013), and that sales competition between teams (stores) only works when the manager and sufficiently many other team members have the same gender. Delfgaauw et al. (2013).

The remainder of this paper is organized as follows. Section 2 presents the experimental design. Section 3 presents the behavioral hypotheses of the experiment. Section 4 comprises the results. Section 5 offers concluding remarks.

2. Experimental design and procedures

2.1. Task description and treatments

Subjects in the experiment are asked to solve a multiplication task, which is commonly used in related studies (see e.g., Kuhnen and Tymula, 2012; Hannan et al., 2013). Specifically, subjects are requested to find the product of a one-digit factor multiplied by a two-digit factor. They do this in five rounds, each round lasting 8 min. After each round, they receive feedback on their performance.

The particular task was chosen for several reasons. It requires no prior knowledge other than basic math skills and it should be easy to understand. Moreover, it is important that performance depends on both ability and effort. This type of task induces heterogeneous ability levels, as math skills are expected to vary greatly; some students study math, others study politics or dance. Students are therefore likely to regard their competence level related to the task differently. Solving math questions is likely to be tiresome, especially when other activities are available in the lab. Specifically, subjects are allowed to engage in two alternatives; read a daily printed newspaper or surf the Internet using their mobile phone.[1] The combination of a simple and tiresome task itself and the alternative activities should induce disutility of effort. Finally, the task provides a stable and precise measure of performance.

Each session has the same sequence of multiplication tasks and all tasks are at about the same difficulty level, thus avoiding any dispersion of results due to variations in the task itself.[2] The screen displays how many minutes are left in each round. Subjects are not allowed to use any type of calculator or any other external remedies. Subjects cannot continue to the next task until they have answered the current task correctly, to avoid strategic behavior of skipping tasks perceived as more difficult. If subjects answer incorrectly, they are told this and asked to try again.

Two dimensions are varied in the experiment, the pay scheme and the type of feedback provided. Subjects are paid either a fixed amount or a piece rate (performance pay) for their participation. The feedback dimension of the design varies across three different performance feedback conditions. In the baseline, subjects receive information on how many tasks they have solved correctly in the previous round. This

[1] A potential worry was that subjects would use their mobile phone to calculate the answers; however, the opportunity to cheat is limited in the lab and easy to detect. Subjects were informed that it was strictly prohibited to use any type of calculator and that they would receive no pay if detected. If someone used a mobile phone as a calculator, the subject would rapidly have to shift focus between the mobile phone and the computer screen, making it easy to detect. I balanced the attention to potential cheating with the concerns of remaining as neutral as possible to avoid any experimenter driven effects (such as subjects feeling monitored or pressured to work hard) (Zizzo, 2010).

[2] I would like to thank Kuhnen and Tymula for sharing the multiplication tasks they used in Kuhnen and Tymula (2012).