Strategic Transmission of Interim Performance Feedback in a Real-Effort Setting: Experimental Evidence

by

Duygu Özdemir

A Thesis Submitted to the
Graduate School of Social Sciences & Humanities
in Partial Fulfillment of the Requirements for
the Degree of

Master of Arts
in
Economics

August 2013

Department of Economics
KOÇ UNIVERSITY

Koc University

Graduate School of Social Sciences and Humanities

This is to certify that I have examined this copy of a master's thesis by

Duygu Özdemir

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____

Asst. Prof. Seda Ertaç

_____

Assoc. Prof. Levent Koçkesen

_____

Asst. Prof. Murat Yılmaz

Date: 13.08.2013

Abstract

In workplaces and schools, interim performance feedback is among the popularly used tools in improving further work motivation and productivity. In some cases and contexts the truthfulness of such feedback can be verified, in others managers or teachers can manipulate the information strategically. In this study, we experimentally analyze the effect of different interim performance feedback mechanisms on the self-confidence of workers, as well as their decisions and future performance, in a "one principal-two agent" context with real effort. We manipulate (1) the verifiability of information, (2) whether the feedback can be given to agents privately or has to be announced publicly. We confirm the theoretical hypothesis that information transmission about the true outcomes of agents occurs only in verifiable feedback mechanisms, including both private and public announcement cases. Yet, agents also show some non-standard behavior, such as not interpreting "no feedback" as negatively as they should, and being influenced by other agents' outcomes although these should be irrelevant. Considering principal behavior, we observe a general tendency to truth-tell, against the theoretical prediction. We also report some results on gender: women are more optimistic than men when strategically given no feedback.

Keywords: Lab experiments, Performance feedback, Strategic information transmission, Cheap-talk, Self-confidence, Lying, Gender

Özet

Ara dönem performans geribildirimi işyerlerinde ve okullarda motivasyon ve üretkenliği artırıcı popüler araçlardan bir tanesidir. Bazı durum ve ortamlarda, geribildirimin doğruluğu kanıtlanabilirken, bazılarında yöneticiler veya öğretmenler bilgiyi stratejik olarak değiştirebilmektedirler. Bu çalışmada, değişik ara dönem performans geribildirim mekanizmalarının çalışanların kararlarına ve ileri dönem performanslarına olduğu kadar, özgüvenlerine olan etkisi deneysel olarak, gerçek eforun kullanıldığı "bir işveren-iki çalışan" durumu üzerinden incelenmektedir. Çalışmada sırasıyla, (1) bilginin doğrulanabilirliği, (2) bilginin kişiye özel olarak veya herkesin gözlemleyebileceği bir şekilde açıklanması durumları üzerinde değişiklik yapılmaktadır. Bulgular gerçek bilgi iletiminin sadece doğrulanabilir geribildirim mekanizmalarında (hem özel, hem de ortak bilgi gönderim durumlarında geçerli olmak üzere) gerçekleştiği hipotezini doğrulamaktadır. Fakat çalışanlar geri bildirim gönderilmeme durumunu değerlendirmeleri gerektiği kadar olumsuz değerlendirmeme ve kendileriyle ilgisiz olmasına rağmen diğer çalışanların sonuçlarından etkilenme gibi bazı standart olmayan davranışlar da sergilemektedirler. İşverenin davranışı incelendiğinde teorik öngörülerin aksine, genel bir doğru söyleme eğilimi gözlemlenmiştir. Stratejik olarak bilgi gönderilmemesi durumunda kadınların erkeklere göre daha iyimser olması gibi cinsiyet bazlı bulgulara da rastlanmıştır.

Anahtar Sözcükler: Laboratuvar deneyleri, Performans geribildirimi, Stratejik bilgi iletimi, Ucuz laf, Özgüven, Yalan söyleme, Cinsiyet

Acknowledgments

I would like to express my deepest gratitude to my supervisor Asst. Prof. Seda Ertaç for patiently providing continuous support, encouragement and exceptional guidance throughout my research. It has been a great experience for me to study under her supervision, which induced me to challenge myself and expand my knowledge. I enjoyed all stages of my thesis with her friendly attitude. I am also indebted to Assoc. Prof. Levent Koçkesen for his persistent guidance and support throughout my thesis. I owe my sincere gratitude for his engagement in the entire process with helpful suggestions and remarks.

I would also thank to my thesis committee member, Asst. Prof. Murat Yılmaz not only for sparing his precious time for participating in my defense and examining my thesis, but also for his valuable comments and suggestions. I wish to thank Prof. Thomas F. Crossley for his guidance and support throughout the analysis process. I would like to thank my professors at Koç University Economics department who enlightened me with their wisdom and knowledge.

Furthermore, I am indebted to the class of MA Economics 2013 for transforming such a hard period to laughter and joy with their special friendship. I am grateful for their support and motivation that gives me courage to follow this path. In addition, I owe a special thank to my dearest friends Zeynep Hazal Özuslu and Müge Şimşek for their genuine friendship with continuous psychological support and encouragement during my graduate study. I thank you all for being a part of my life.

Most of all, I would like to thank my family for their continuous love and support throughout my entire life. I thank them for encouraging me to pursue my dreams.

Lastly, I am grateful to TÜBİTAK(through Dr. Ertaç's grant 111K444) and Koç University Graduate School of Social Sciences and Humanities for their financial support for conducting the experiments, and Vehbi Koç Foundation for their financial assistance granted throughout my graduate study.

Contents

List of Figures

List of Tables

To my loving family...

1 Introduction

In workplaces and schools, work motivation is the key element to increase the performance and success of individuals. Interim performance feedback is among the popularly used tools in improving further work motivation and productivity. While firms use monthly or annual performance evaluation as a performance feedback, schools and universities employ grade revelation for the same purpose. According to the evidence cited by Murphy and Cleveland (1991), around 80 % of the firms exercise a formal interim performance feedback system as an evaluation technique.

While feedback is an important potential motivational tool in many settings, different contexts bring different constraints in terms of how information can be revealed. In some settings and contexts the truthfulness of performance feedback can be verified (e.g. all information that is revealed should be truthful), in others managers or teachers can manipulate or withhold the information strategically. Likewise, there can be differences in terms of "audience". Sometimes feedback can be given in private, whereas sometimes the performances can be observed publicly by all of the workers or students. However, which one of those mechanisms should be preferred in a performance-based environment regarding efficiency concerns?

In this paper, we experimentally evaluate alternative interim performance feedback mechanisms in terms of their impact on equilibrium information transmission, as well as their effect on the self-confidence of workers and future performance. We use a “one principal-two agent” context with real effort, where agents work for (potentially) two periods. Whether to work in the second period is a choice of the agent, and since high performance by agents in the second period benefits the principal, the principal wants both agents to decide to work, and then perform well. The principal observes the first period performance of agents and then decides what to reveal to the agents. We manipulate in five treatments (1) the verifiability of information, (2) whether the feedback can be given to agents

privately or has to be announced publicly.

We confirm the theoretical hypothesis that information transmission about the true outcomes of agents occurs only in verifiable feedback mechanisms, including both private and public announcement cases. We also find that verifiable reports in a multiple-audience context, i.e. publicly announcing, is the most efficient mechanism which yields the highest future willingness to perform and has a significantly positive effect on performance controlling for the type of feedback.

We use individual performance revelation in which outcomes of agents are independent, and thereby we aim to eliminate the social comparison concerns on performance. Although a rational agent is theoretically expected to be insensitive to other agent's message in this context, we observe that agents are negatively influenced by the success of the other agent in public cases which proves the existence of psychological factors affecting the decision-making process.

In terms of the principal's strategy, we find that under bad outcome in cheap-talk cases, where lying is monetarily incentivized, subjects prefer rather to reveal the true information or to transmit no information. Therefore, there is a general tendency to truth-telling, against the theoretical prediction. Thus, regarding the similar results of the previous experimental studies in different contexts (Gneezy (2005), Charness and Dufwenberg (2006), Abeler et al. (2012), Fischbacher and Föllmi-Heusi (2013)), we conclude that honesty is behaviorally a general characteristic of a sender-receiver game. Also considering the results of agent-side, agents underestimate the propensity of principal to post truthful feedback. We additionally provide a novel insight to reporting strategy by exploring the expectations of the principal. We analyze the truth-telling result via principal's expectations and find a significant negative impact on the realization rate of a bad outcome. In other words, as principal expects agent to be more self-confident, they are less likely to reveal the true failure result. This is the first study that collects data for expectations of the feedback sending authorities which is explanatory about the

decisions in message sending process.

Our paper brings together several strands of the literature. While the effects of performance feedback on beliefs and effort has been a recently popular topic of research (Ederer and Fehr (2007), Rosaz (2012), Gürtler and Harbring (2010)), there is no study that explores strategic information transmission with two audiences, in a real-effort context. To our knowledge, this is the first comprehensive study that explores information transmission under alternative mechanisms in terms of audience and verifiability, and exploits a real-effort setting to uncover the impacts of performance feedback realistically.

The rest of the paper proceeds as follows: section 2 summarizes the related literature, section 3 reviews the design of our experiment, section 4 explains the procedures and the nature of our data, section 5 exhibits our theoretical model and the hypotheses, section 6 displays our results and discussion, and section 7 concludes.

2 Related Literature

Our paper is related to a number of strands of the literature. In psychology, management and organizational behavior literatures, performance feedback has long attracted researchers' interest as a motivational tool (Barr and Conlon (1994), Gibbs (1991), Murphy and Cleveland (1991)). However, it is a relatively recent topic in economics. Theoretical studies on feedback revelation generally employ principal-agent models to analyze the optimality of the feedback mechanisms (Ertac (2005), Ederer (2010), Aoyagi (2010)). In laboratory experiments, the effects of the feedback on self-confidence, motivation and further performance are investigated. For example, Ertac (2011) analyzes the question how self-relevant performance feedback influences beliefs and self-confidence. It points that under negative feedback, subjects with higher initial self-confidence are more optimistic, where initially unconfident subjects are more pessimistic in belief updating. Another study, Eil and Rao (2011) shows that people are more likely to update their beliefs under positive feedback than negative in a relative-performance setting. In addition to the asymmetry in belief updating, Mobius et al. (2011) also finds conservatism such that people little respond to both positive and negative signals when the context is ego-relevant.

Beside these studies which focus more on belief and self-confidence, there are a number of studies that explore the effects of feedback on future performance and motivation. Azmat and Iriberri (2010) analyzes the theory on a natural experiment data and find a strong positive effect of relative performance feedback on effort. On the other hand, Gill and Prowse (2012) observes a drop in motivation due to the negative feedback.

All of the aforementioned studies concentrate on relative performance feedback, which is automatically and truthfully revealed. Yet, in reality, performance feedback is mostly announced by a principal, i.e. a director or a teacher, and can be manipulated or withheld, in order to motivate workers or students and lead them

to enhance further performance. The idea of strategic information transmission by a principal is a relatively recent topic in economics.

In the economics literature, such reporting mechanisms are extensively studied in various contexts. Theoretically, the pioneering works such as Grossman (1981) and Milgrom (1981), have analyzed verifiable information in a signaling game in which they find a full-revelation equilibrium if the sender has monotonic preferences. The sufficient conditions for that result were updated according to a more general context by Seidmann and Winter (1997). On another strand of feedback mechanisms, Crawford and Sobel (1982) investigates unverifiable information, i.e. cheap-talk, where sender was allowed to lie arbitrarily. These studies inspired several further works both theoretically and experimentally. Regarding multiple-audiences, Koessler (2008) and Farrell and Gibbons (1989) explore the verifiable and cheap-talk reporting theoretically in a one sender-two receiver setting, respectively. While Koessler (2008) exhibits mutual subversion as the equilibrium strategy in verifiable-public case, Farrell and Gibbons (1989) asserts mutual discipline as the possible strategy in equilibrium of cheap-talk public[1]. Following the theoretical works, Battaglini and Makarov (2010) verifies experimentally the results of Farrell and Gibbons' theory.

Note that all of the above discussed studies that focus on verifiable and unverifiable reports are about revealing the sender's type and do not consider any context of signaling the type of receiver. There are very few studies that are more relevant to our context, in which the report is associated with the receiver's type.

Regarding cheap-talk mechanisms in principal-agent contexts, Ederer and Fehr (2007) provides an important result with a one principal-two agent tournament setting. They observe that principals report the difference in performances less than it actually is. Also, in contrast to their predictions, the message of principal

[1]Mutual discipline refers to the no information revelation to any of the receivers, but a fully revealing equilibrium is the case when the public communication employed. On contrary, mutual subversion denotes the case where full-revelation equilibrium occurs in private case while information is not revealed when communication is public.

is informative and has a negative impact on agent's effort. Another research by Rosaz (2012) displays that principal positively manipulates the message.

In addition to cheap-talk, Gürtler and Harbring (2010) theoretically and experimentally investigates relative performance feedback in verifiable reporting mechanism. In a tournament-based setting, agents' performances are determined by randomly appointed parameters. Their results suggest a tendency of principals to reveal the small gaps in performances of agents, while principals are more likely to withhold the bigger performance gaps. Unlike the theoretical predictions, principals withhold the information more frequently than expected.

However, none of these papers focus on a real-effort setting, and rather induce ability and effort hypothetically. Although this type of design is advantageous in terms of controllability of unobservables and simplicity of its application, their inference about such concepts as self-confidence, motivation and performance is unrealistic. Creating an ego-relevant environment through a real-effort task is important for validity and applicability of the results to real-life settings.

We also report some results related to gender in our paper. There are a few existing studies on the interaction of gender and the interpretation of performance feedback. Ertac and Szentes (2010) analyzes the degree of gender difference in competitiveness through performance feedback. Their results suggest that the gender based competitive characteristics converge to each other with the provision of the feedback. Also Barankay (2011) aims to uncover the effects of relative feedback on performance by a field experiment. He detects a demoralizing effect of the negative feedback which is only influential on male participants. On the other hand, Gill and Prowse (2010) observes a drop in motivation of female after losing in a tournament. Yet, we do not use tournament-based setting in our study. Our study contributes to this strand of the literature by exploiting an individual setting to analyze the gender effect on both principal and agent domain. Our results suggest that women are more optimistic than men when strategically given

no feedback. On the principal's side, we do not find any significant correlation between gender and the reporting behavior.

Finally, standard economic theory assumes a zero lying cost such that people lie when it is a payoff maximizing option. However, recent empirical studies show that honesty and truth-telling is also a preference even under unverifiable contexts (Kartik (2009) and supported by the experimental study Sanchez-Pages and Vorsatz (2009)). According to several studies, subjects face a psychological lying cost when they misreport their type (Kartik et al. (2007), Chen et al. (2008), Kartik (2009)). In the intersection of the results of several experimental studies, misreporting the private info is a frequently chosen option, yet a substantial share of subjects avoid lying even though it is incentivized monetarily(Gneezy (2005), Charness and Dufwenberg (2006), Fischbacher and Föllmi-Heusi (2013) and Abeler et al. (2012)). Although these results are obtained in different settings, they are robust in our context, as well.

3 Experimental Design

The experimental design is based on studying interim performance feedback in a "one principal-two agent" real effort context. The performance feedback transmission mechanism is the treatment, whereby the performance revelation technology available to the principal changes between periods. We explore the effects of 5 different treatments in a within-subject design. The experiment consists of 5 periods and within each period there are two rounds. To eliminate potential wealth effect, we use a random payment scheme, i.e., one of the 10 rounds is chosen randomly and subjects are paid according to their payoffs in that chosen rounds.

At the beginning of the experiment, subjects are randomly appointed as either "Principal" or "Agent". In each period, 3-person groups, which consist of 1 principal and 2 agents, are formed. In group forming process, stranger design[2] is used, in order to prevent reciprocity and revenge motivations of participants.

For agent type participants, we employ two types of tasks: an addition task and a verbal task[3]. The verbal task consists of general knowledge questions, as well as, verbal classification and number-letter matching questions. The addition task involves adding 4 or 5 two-digit numbers. The purpose of using different tasks is to distinguish the effects of the feedback on different types of performance.

In each period, agent type subjects are randomly assigned to one of these tasks and perform the same task in both rounds of that period. For both tasks, subjects are asked to solve as many questions as possible within a limited time (120 sec.). At the end of each round, the number of correct answers, the "score[4]", is compared to the target score[5] chosen for that specific period. The appointed target score is employed in both rounds of the corresponding period. If a subject's score is greater than or equal to the target score, the subject is "Successful". The subject

[2]In stranger design, subjects are re-matched randomly at the beginning of each session.

[3]We provide details of the applied tasks in Appendix 1. Instructions part.

[4]Note that wrong answers does not affect the score.

[5]Target score is a number which is randomly chosen from the interval [4,13] at the beginning of each period. The interval of the target number is determined with respect to the data from pilot study.

has "Failed" otherwise. Note that the target score is subject-specific and there is no common shock applied to the performance of subjects.

3.1 Belief elicitation

To elicit self-confidence, we use a crossover mechanism as in Mobius et al. (2011)[6]. In this mechanism, subjects are provided two lotteries to determine their second-round payoff. At the first lottery, they achieve the reward if the outcome is "Successful" in second-round performance stage. At the second lottery, they earn the reward with a probability X, regardless of their performances. At the end of the first performance round, subjects are asked to report the minimum probability of chance, X, to choose the second lottery. Here, a rational subject is expected to report the subjective probability she assigns to being "Successful" for the second round regardless of her risk preferences. We ask the subjects to make this decision before and after feedback to see the confidence level prior with pre-feedback report as well as how and at what magnitude they update their beliefs of success for second round after receiving information with post-feedback report. To address incentive-compatibility concerns, one of the stated beliefs is randomly chosen and either performance or chance mechanism is applied. Thus the subjects are incentivized to state their true beliefs for both pre- and post-feedback reports.

The timeline of a period for agents is as follows:

1. Pre-feedback performance: Subjects perform the appointed task within 120 sec.

2. Pre-feedback belief: Without receiving any message, subjects state the minimum probability of chance to choose the second lottery.

3. Feedback: The feedback type is changing between periods which will be explained in next part.

[6]The crossover mechanism is a type of a Becker-DeGroot-Marschak(BDM) mechanism as described in Mobius et al. (2011).

4. Post-feedback belief: The subjects are allowed to update their previous report.[7]

5. Performance/Chance mechanism: If the randomly chosen report is higher than the probability determined by chance, then the subject performs the same type of the task in 120 sec. as in first round. Otherwise, they do not perform and their second-round payoff is determined by chance.

3.2 Feedback mechanism

Note that after the first round, only the principal is informed about whether the agent has been successful or not. Agents have no information about either self-score or the appointed target score. Since they solve the questions, they have prior beliefs about their self-score. With the feedback they receive, they infer whether they successfully pass the target score. There are three types of feedback mechanisms used throughout the experiment:

1. Truthful feedback mechanism: In this mechanism, subjects receive an accurate message from the computer. This is the baseline mechanism in our design.

2. Verifiable feedback mechanism: The interim feedback is reported by the principal in this mechanism. Here, principal chooses either to transmit the true outcomes of agents, or to withhold the information. There is no deception in sent messages since they must be correct.

3. Unverifiable, i.e. Cheap-Talk, feedback mechanism: As in the verifiable mechanism, principal sends the feedback to the agents. However, the prin-

[7] At this stage, the subjects can see their previous reports on the screen.

cipal does not have to report the actual outcome in this mechanism. Regardless of the outcomes, the principal has "Successful", "Failed" and "No Information" message options that allows her to lie. Also the mechanism involves withholding the information through which we intend to capture behaviorally whether there is a non-zero lying cost that would lead principals to refrain from lying.

Within these feedback mechanisms, except the baseline, we also employ two different types of mechanisms that differ in the audience of the messages:

- Private feedback: In this mechanism, principal reports the feedback independently and privately to agent and agents only see the message targeted to them.
- Public feedback: Unlike private, principal has to announce the feedback publicly. In this mechanism, agent observes the other agent's feedback, in addition to her own feedback. The options provided to principals in public case vary treatment-specifically. In the verifiable-public case, principal decides either to release the truthful outcome to both of the agents publicly, or to withhold the information. On the other hand, in cheap talk case, the feedback for each agent is chosen separately from the three options explained above and feedbacks for both agents are commonly reported.

In addition to feedback report, principals are asked to guess the reported minimum probability of chance (X) to prefer chance mechanism after feedback by each agent. Guessing process is incentivized by paying 10 ECU for each correct[8] estimate.

[8]If the guessed amount is in the $\pm$ 5 interval of agent's actual decision, the estimate would be regarded as correct.

3.3 Payoff formation

3.3.1 Agent payoff

The payoff of agent type participants is formed according to the outcome of her performance and her decisions. To preserve the salience of actions and incentivize the performance, we employ a discrete payoff which involves a fair amount of monetary difference between the outcomes. Here, the payoff of agent in performance mechanism is determined as following:

300 ECU, 18 TL[9], if the agent is "Successful",

100 ECU, 6 TL, if "Failed".

So the earnings of first round is formed according to the performance payoff mechanism. For second round, if the agent prefers performance, the payoff again depends on the performance outcome. However, if the agent chooses chance mechanism, then earning of the second round is 300 ECU with probability X, and 100 ECU with probability (100-X) where X is the exogenously appointed probability of chance.

3.3.2 Principal's payoff

In our design, we employ a utility function where the performances of both agents are regarded as complements. The payoff function of principal is:

$$U_p^t = \begin{cases} 100 & , t = 2n-1 \\ 50 + 10(g_t^1 + g_t^2) + min\{q_1, q_2\} & , t = 2n \end{cases}$$

where

q_i: performance return of agent i, $q_i(c_{it}, e_{it}) = \begin{cases} 20 * c_{it} & , e_{it} = 1 \\ 0 & , e_{it} = 0 \end{cases}$

g_t^i: guess at round t for agent i, where 1 if guess is correct, and 0 otherwise

n: period where n ϵ [1,5]

[9]Here, the amounts presented is the equivalent amount of the stakes in Turkish Liras.

Note that c_{it} denotes the number of correct answers of agent i, at round t, while e_{it} represents the entry of agent i, at round t, to the performance stage. Here, at the first round the principal's payoff is a constant amount, 100 ECU. The second round payoff is composed of three elements; a constant amount, 50 ECU, extra 10 ECU for each correct guess and the minimum of the returns from both agents. As presented above, for a positive return besides the constant earning, agent must enter the performance stage. Since the function includes minimum condition, both of the agents have to prefer performance mechanism for second round to provide principal a non-zero payoff from the part of function that depends on agent performance. Thus, the aim of the principal is not only to convince both agents to prefer performance mechanism, but also to maximize the post-feedback performances of both agents via feedback.

4 Procedures

The experiment was programmed using z-Tree experimental software (Fischbacher (2007)). It was conducted at the Koç University and Bilgi University computer labs in the spring term of 2013. We collected data from 132 subjects in total (72 subjects from Koç University and 60 subjects from Bilgi University). The experiment was held in 13 sessions, with 8 sessions at Koç University and 5 sessions at Bilgi University. At the end of each session, we conducted a survey to collect demographic and personality data such as age, gender, faculty of the subject and GPA. Also to eliminate the order effect of treatments, we used 6 different configurations[10]. Sessions lasted about 50 minutes, and subjects earned on average 20,23 TL, 10 TL of which was a show-up fee[11].

Our subject pool consists of 68 male and 64 female participants. The majority of the subjects are undergraduate students registered from diverse faculties[12], yet there are a few graduate students[13], as well.

Before the actual data collecting sessions, a pilot session was conducted to check whether the instructions are easily comprehensible. Also some parameters of the experiment such as stakes, exchange rate and the interval of target score are determined according to the data collected in pilot.

[10]The configurations are recorded as follows: 1 for TVC, 2 for TCV, 3 for VCT, 4 for CVT, 5 for VTC and 6 for CTV. Note that T, V and C corresponds to Truthful, Verifiable and Cheap-Talk feedback mechanisms respectively. In Verifiable and Cheap contexts, we used private and public sequentially without changing the order.

[11]Earnings ranged between 15 TL and 28 TL, including the show-up fee.

[12]There are at least 3 subjects from each faculty. Though, nearly half of the subjects are students of CASE faculty.

[13]There are 9 graduate students over 132 subjects who participated in the experiment

5 Hypotheses

In our design, the state of the world θ_i is binary that is either a high performance, h, or a low performance, l. Principal reports $r_i \in M_i = \{l, h, \emptyset\}$, where $\emptyset$ represents no report case. Principal's strategy is a pair $\rho = (\rho_1, \rho_2)$ with $\rho_i : \{l, h\} \to \{l, h, \emptyset\}$, $i = 1, 2$. With the principal's report r_i, agent i forms beliefs about θ_i which we denote $\mu_i : \{l, h, \emptyset\} \to [0, 1]$. Here, $\mu_i(r_i)$ is the probability that agent i assigns to the $\theta_i = h$ event. As we explained in the experimental design section, we employ four different report mechanisms; two of them differs in terms of accuracy of feedback, Verifiable and Unverifiable, i.e. Cheap-Talk, feedback, and the other two of them varies in the audience of the feedback, Private and Public feedback. These give four different games, in which we formed our hypotheses below.

1. In Verifiable-Private mechanism, full information revelation is the unique PBE, i.e. in equilibrium $\rho_i(h) \neq \rho_i(l)$ for i= 1, 2.

2. In Verifiable-Public mechanism, full information revelation is always an equilibrium. However, under certain conditions, partial information transmission is also a PBE, such that only (h,h) is revealed.

3. In Cheap-Private mechanism, there is no credible feedback in equilibrium such that in any PBE, $\rho_i(h) = \rho_i(l)$, i= 1, 2.

4. In Cheap-Public mechanism, no information transmission is always an equilibrium. Yet, partial information transmission is also a PBE under certain conditions.

5. In public feedback case, since the tasks and target scores are independent across agents, the feedback of other agent should theoretically be irrelevant in belief updating. That is only related to the strategy of the principal. Therefore, we expect the belief updating to be independent of the feedback about the other agent. If not, there are psychological factors that subjects

incur in public cases.

6. According to Bayes' Rule, agents are expected not to update their beliefs under "no information" case.

7. In theory, the cost of lying is assumed to be zero since a rational agent is expected to lie under bad outcome condition to maximize her payoff. However, with the "No information" message option in cheap-talk case, we intend to capture whether there is any non-zero lying cost.

6 Results and Discussion

We conduct our analyses from two perspectives; agent and principal side. In agent side analysis, we investigate whether information transmission occurred and how agents interpreted the message in updating their beliefs. In principal side analysis, we explore the strategy of principals in message sending process. For both approaches, we also investigate whether there is a gender difference. Results of agent and principal side analyses are discussed in next two subsections respectively.

6.1 Agent Side Analysis

Before delving deeper into the analysis, we provide some summary statistics about task performance. On average (in both rounds), subjects attempted to solve 9.3 questions in addition task and 11.2 questions in verbal tasks. They correctly solved 7.04 and 7.85 questions on average respectively. Post-experiment survey results show that subjects evaluate the difficulty of both tasks as moderate[14]. The answers to a survey question which asks whether it is important for subjects to succeed independently of its monetary payoff reveal that a majority of subjects do care about success, which means that we have managed to create an ego-relevant performance environment for subjects in our experiment[15].

We firstly examine prior confidence level composition of the agent type subjects. Pre-feedback beliefs of agents exhibit that agents mostly prefer to perform (mean= 66.02%, median= 70%). Since there is a strong tendency to enter performance stage and only 51% of the entered subjects successfully passed the target score, we conclude that participants overestimate their performance. It is observable in Figure 1, and is verified by Mann-Whitney test, that female participants are less confident, likewise in the literature (z= 4.494, p= 0.0000). We also check for a task-based difference in confidence levels. In Table 1, the average and me-

[14]On average, subjects assessed 2.86 for Addition and 2.87 for Verbal task. For more information see Appendix 3, questions 7 and 8.

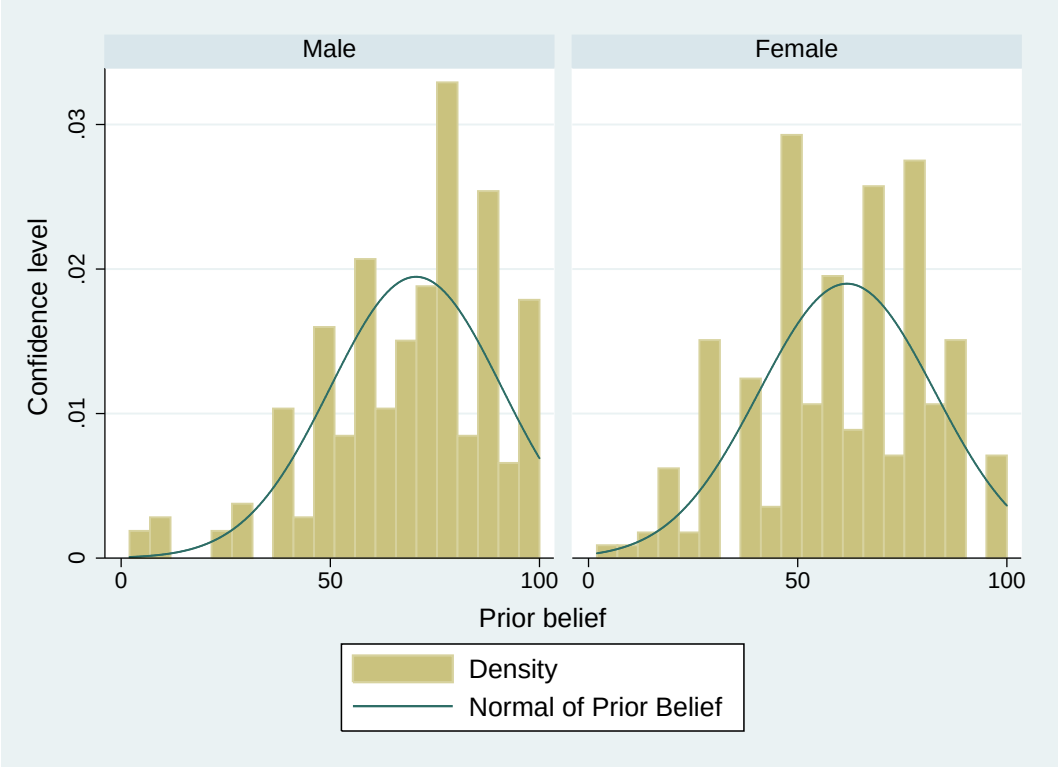
[15]The mean assessment of subjects is 3.74, where 75.82% of them chose either important or very important choice. For more information see question 10 in Appendix 3.

dian stated pre-feedback beliefs are presented. We observe that subjects are less confident in verbal task (z= -1.962, p= 0.0497 according to the Mann-Whitney test).

Table 1: Confidence level by Tasks

	Mean(%)	Std. Dev.	Median(%)
Addition task	67.76	21.74	70
Verbal task	64.26	20.53	65

Figure 1: Comparison of Prior Confidence Level Distributions by Gender



Besides prior confidence levels, we investigate the performance in both types of tasks. Even though there is no substantial gap between mean and median scores in both tasks(mean score is 6.72 in Addition task and 7.43 in Verbal task), Mann-Whitney test points out that subjects have lower initial performances in addition task (z= 2.832, p= 0.0046). Therefore, while subjects perform worse in the first round of the addition task, they state higher pre-feedback belief. After

the performance in addition task, subjects more or less have an idea about their scores whereas the score is more vague in verbal task since it requires knowledge and ability. Therefore, the addition task is likely to be a much less ambiguous task that probably leads subjects to be initially more confident.

As a starting point for the agent side analyses, we explore whether prior beliefs are independent of treatments and configurations applied throughout the experiment to ensure that there is no bias in our further analyses. In regression analysis, we do not find any significant treatment or configuration effect on pre-feedback beliefs which confirms the robustness of prior beliefs overall(see Table 2)[16]. This makes us able to assume that beliefs are initially same across all treatments, and configurations. We also check for a difference between the prior beliefs, the confidence levels, of Bilgi and Koç students. Mann-Whitney test certifies that there is no statistically significant difference at all (z= 1.010, p= 0.3125). This rules out the possibility of school effect on confidence characteristics of the subject pool.

[16]Note that we collect data from the same subject in all of the treatments. Thus, to account for correlation, we use random effects model in the regressions require multiple observations from the same subject.

Table 2: Treatment and Configuration Effect on Initial Belief

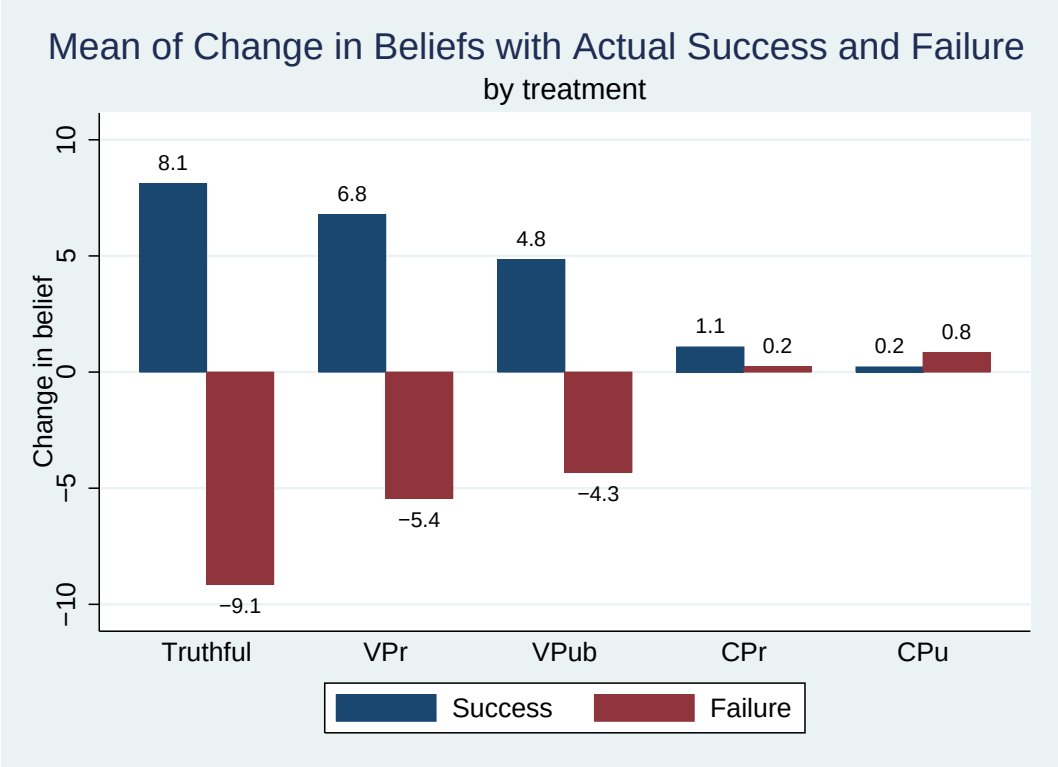
	(1)	(2)	(3)
Truthful(1,0)	3.329		3.358
	(2.270)		(2.274)
Verifiable-Private(1,0)	2.022		2.050
	(2.270)		(2.274)
Verifiable-Public(1,0)	2.351		2.380
	(2.270)		(2.274)
Cheap-Private(1,0)	2.978		3.006
	(2.270)		(2.274)
Cheap-Public(1,0)	.		.
	.		.
TVC(1,0)		0.699	0.522
		(5.791)	(5.764)
TCV(1,0)		-1.556	-1.556
		(6.717)	(6.684)
VCT(1,0)		8.563	8.563
		(6.463)	(6.431)
CVT(1,0)		2.677	2.677
		(6.358)	(6.327)
VTC(1,0)		3.000	3.000
		(6.581)	(6.549)
CTV(1,0)		.	.
		.	.
N	447	447	447
χ^2	2.574	3.069	5.712

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

GLS Regressions for Different Feedback Mechanisms

Figure 2: Information transmission under different treatments



The main focus of our study is information transmission to agents in equilibrium across different treatments. In order to analyze this, we investigate the effect of actual success/failure outcome of the agent (which is always observed by the principal but is unobservable to the agent except in the truthful case) on her belief updating process. Figure 2 displays the varying intensity of information transmission between treatments, in which the ordinal sequence of the change in beliefs is from truthful as the most to cheap-public as the least transmission occurred. We also estimate Ordinary Least Squares (OLS) regressions with change in belief as the dependent variable to check the significance of transmissions. According to the results shown in Table 3, the information about the actual outcome has a significant effect on agents' belief updating decisions in truthful and verifiable feedback cases, while it has no significant impact in cheap-talk case. This result reveals that information about the true outcome is transmitted only in truthful and

verifiable feedback mechanisms, which agrees with our theoretical predictions. In other words, the true information is not transmitted in cheap-talk case. Whether this comes from the agent's updating or the principal's strategy will be explored in the next subsection.

Table 3: Effect of actual success on change in belief under different treatments

	(1)	(2)	(3)	(4)	(5)
	Truthful(Baseline)	Verifiable-Private	Verifiable-Public	Cheap-Private	Cheap-Public
success	17.74***	12.28***	8.141*	-0.302	-1.282
	(5.523)	(3.787)	(4.420)	(3.217)	(3.447)
session	YES	YES	YES	YES	YES
N	91	91	91	91	83
R^2	0.217	0.230	0.214	0.0774	0.0786

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

OLS Regressions for Different Feedback Mechanisms

Another core point of our analyses is the effectiveness of messages, which depends on both principal's strategy and agent's perception. Since we will go through the principal's strategy in the next subsection, we now analyze how agents interpreted messages they received. In Figure 3, Figure 4 and Figure 5, the impact of feedback on belief updating in different feedback mechanisms under positive, negative and no feedback cases is displayed, respectively. A Wilcoxon signed-rank test shows that both positive and negative feedbacks have significant effects in all treatments[17], except for the cheap-private case (p= 0.2359 for positive feedback and p= 0.4871 for negative feedback case). On the other hand, it is confirmed by a Wilcoxon signed-rank test that "no feedback" has no significant impact on belief updating in any of the treatments (p= 0.2109, p= 0.9804, p= 0.6741, p= 0.7098 respectively for VPr, VPu, CPr and CPu). To scrutinize the effect deeper, we perform OLS regressions of change in belief on each type of feedback with some additional controls. Here in Table 4, we find that positive feedback has a

[17]The results of Wilcoxon signed rank test for each treatment is as following: truthful with p= 0.0002 and p= 0.0002, VPr with p= 0.0002 and p= 0.0466, VPub with p= 0.0021 and p= 0.0033, CPub with p= 0.0050 and p= 0.09 for positive and negative feedback cases respectively.

significant positive effect on belief updating across all treatments, except cheap-private, consistently with the result from the non-parametric test. This implies that people consider cheap-public feedback as credible, whereas they do not treat cheap-private information as reliable as in the public case. Similarly, we show that there is no significant impact of "no feedback" on belief updating under any feedback mechanism[18].

We then extend our analyses to question whether the message about the other agent affects belief updating in public cases. As it is seen in column (4) and column (7) of Table 4, controlling for one's own performance, observing the success of another agent has a significant negative effect on the improvement between pre- and post-feedback beliefs. Thus, even though subjects are instructed that the outcomes and messages are independent in terms of both type of tasks and applied target scores, they are psychologically influenced by the message about other agent.

[18] Analyzing the effects of all "true feedback" (coming from Truthful, VPr and VPu treatments) again shows that true positive feedback significantly improves the change in beliefs, and there are no differences in treatment, as long as the feedback is truthful.

Figure 3: Impact of Positive Feedback on Belief Updating

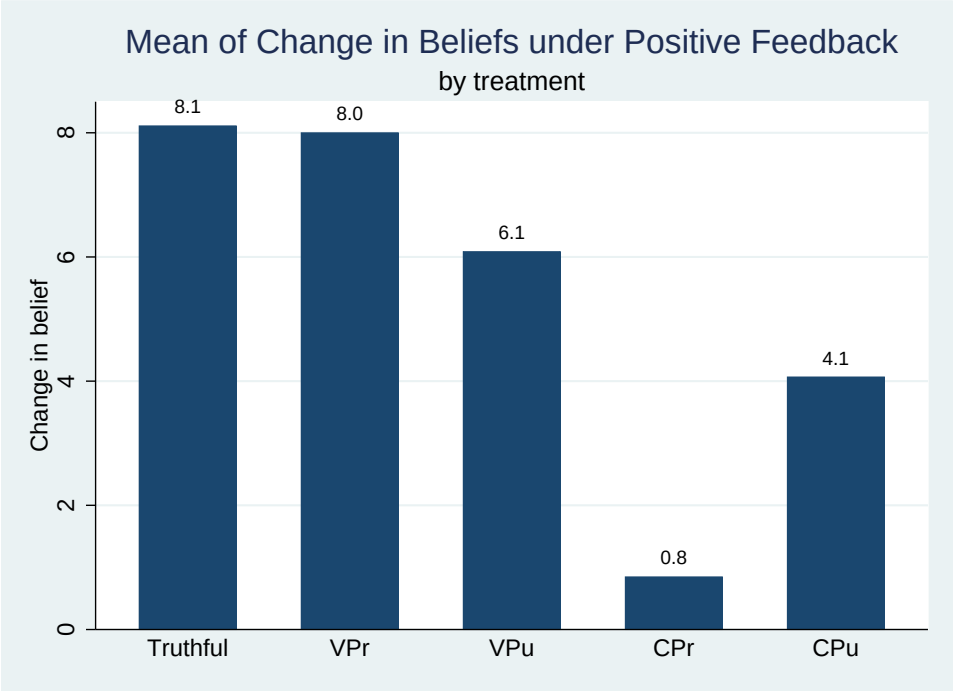


Figure 4: Impact of Negative Feedback on Belief Updating

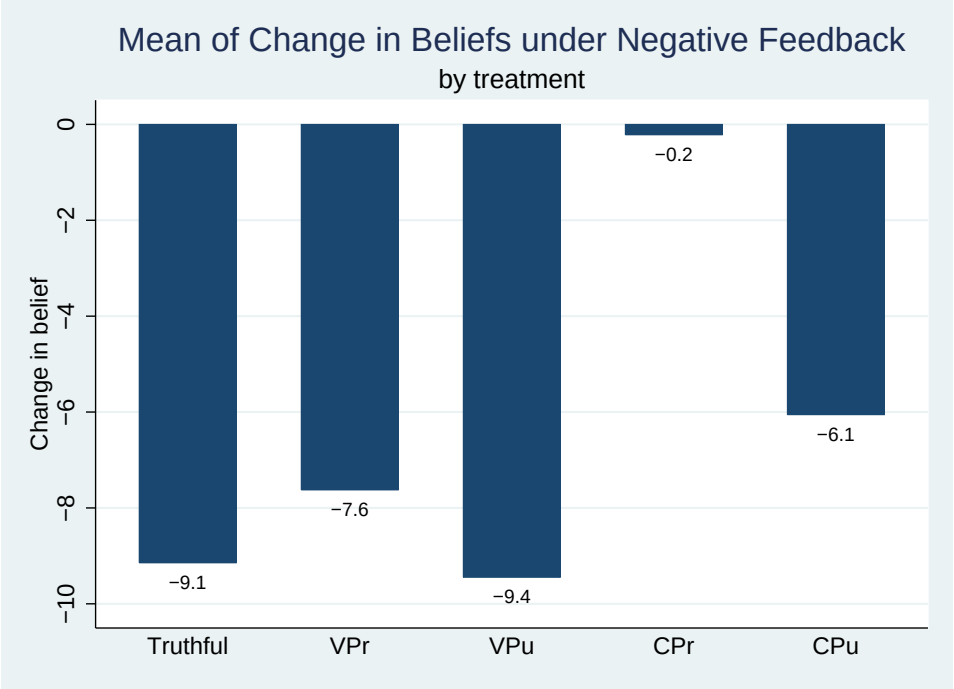


Figure 5: Impact of No feedback on Belief updating under different treatments

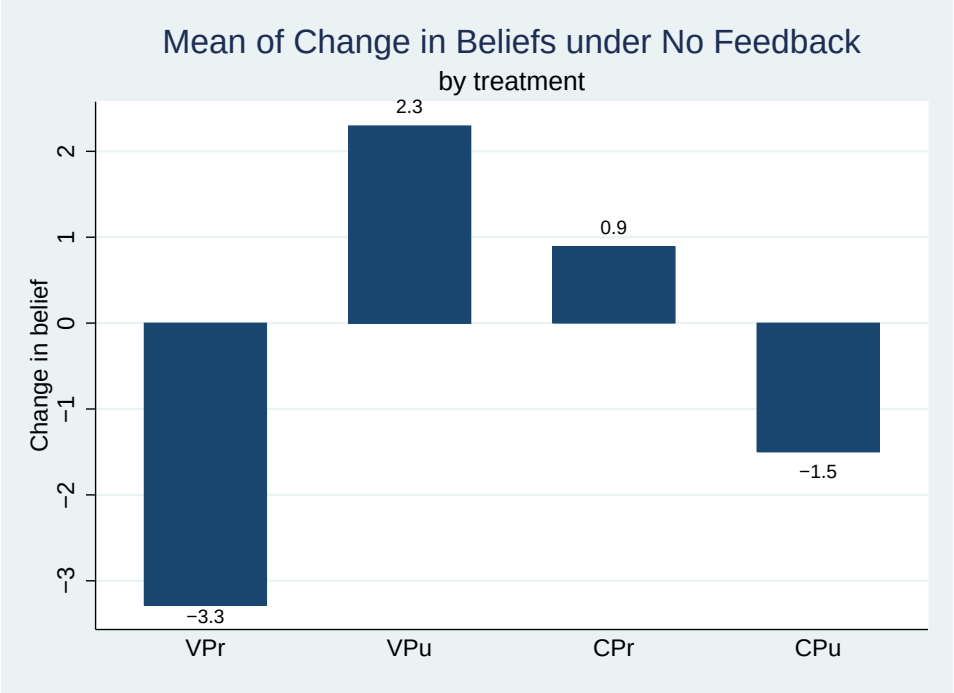


Table 4: Effect of Feedback on Belief Updating

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Truthful(Baseline)	Verifiable-Private	Verifiable-Public	Verifiable-Public	Cheap-Private	Cheap-Public	Cheap-Public
Positive Feedback(1,0)	18.31***	15.48***	12.57**	10.70**	-1.765	12.91***	13.32***
	(5.594)	(4.796)	(5.189)	(5.208)	(4.263)	(4.365)	(4.287)
No Feedback(1,0)		2.467	8.636	2.360	-0.517	6.680	6.610
		(4.944)	(5.998)	(6.820)	(6.613)	(5.070)	(5.052)
Female	7.059	7.722*	-4.614	-4.833	-0.466	-3.118	-2.179
	(5.045)	(4.078)	(4.741)	(4.668)	(3.430)	(3.453)	(3.416)
Period	-1.839	-2.416	7.461	7.590	0.668	0.437	-0.275
	(3.055)	(4.045)	(4.763)	(4.688)	(2.955)	(3.124)	(3.086)
Addition task(1,0)	1.199	-5.471	-8.564**	-7.959*	0.326	-5.234*	-5.618*
	(4.465)	(3.442)	(4.150)	(4.098)	(3.071)	(3.027)	(2.977)
Session	YES	YES	YES	YES	YES	YES	YES
Positive feedback				-9.563*			-9.027**
of other agent				(5.201)			(4.239)
No info							-7.459
of other agent							(4.738)
N	91	91	91	91	91	83	83
R^2	0.237	0.283	0.298	0.329	0.0803	0.219	0.270

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

OLS Regressions for Different Feedback Mechanisms

Besides feedback, we explore how other factors affect the belief updating process under certain types of feedback. We perform Generalized Least Squares (GLS) regressions of the change in beliefs on several factors, with controlling for treatments. In Table 5, we find that addition task has a significant negative impact only when the subject receives a positive feedback. This implies that under positive feedback, the magnitude of the change in belief is significantly less in the addition task than in the verbal task. Also we observe in the same table that period has no impact on change in belief under any types of feedback which indicates the robustness of the belief updating against the timing of decision and experience.

Table 5: Gender difference on message interpretation

	(1)	(2)	(3)	(4)
	Positive feedback	Non-positive feedback	Negative Feedback	No Feedback
Female	-3.334	5.249**	4.472	5.106*
	(2.794)	(2.471)	(3.728)	(3.078)
Addition task(1,0)	-4.225**	-2.266	-2.833	-0.613
	(1.643)	(2.446)	(3.508)	(2.850)
Period	0.661	-0.619	-0.907	1.219
	(0.647)	(0.924)	(1.264)	(1.410)
Verifiable-private(1,0)	-1.838	3.899	1.285	-0.0432
	(2.885)	(3.543)	(5.205)	(5.467)
Verifiable-public(1,0)	-5.231*	5.918	0.894	3.350
	(2.798)	(3.656)	(5.031)	(5.260)
Cheap-private(1,0)	-8.350***	9.885**	9.395*	0
	(2.544)	(4.206)	(5.383)	(0)
Cheap-public(1,0)	-6.268**	6.486	4.742	-0.401
	(2.663)	(4.057)	(5.837)	(5.482)
N	207	240	149	91
χ^2	24.78	11.58	6.594	.

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

GLS Regressions for Different Feedback Mechanisms

Yet, we also add a gender-based approach to our analysis portfolio. With the results presented in the column (2) of Table 5, in non-positive feedback condition[19], female has a positive effect on belief updating. Therefore, we conclude

[19]By non-positive feedback, we consider either no or negative feedback transfer to the agents.

that women update their beliefs more optimistically under non-positive feedback cases. In order to clarify that result, we examine non-positive feedback separately as negative and no feedback, as well. By the results presented in column (3) and (4) on the same table, we observe that female participants significantly increase their beliefs 5% more than male counterparts when they receive no feedback. This indicates that women do not regard "no feedback" as a negative signal and their optimism in belief updating is specifically valid in an ambiguous state rather than a negative state.

As an alternative analysis, rather than change in belief, we perform the same regressions on posterior beliefs by controlling prior beliefs. Our aim is to check whether the results are consistent if we take dependent variable to be the posterior belief rather than the magnitude of change in beliefs. The results in all regressions are robust under that alternative model.

Besides the analysis of belief-updating, we examine the influence of interim feedback on performances and the tendency to work in the second round. Here, we question which one of the feedback mechanisms has the most efficient outcomes in performance-based environments. First, we investigate the pre- and post-feedback performances across treatments. Note that in our design, subjects are allowed not to choose to perform in the second round. Thus, for the validity of comparison, we considered initial performances at the level of subjects who ended up performing in the second round and at the aggregate level (see Figure 6). We also survey the entry rate to the second round performance stage, in treatment level (see Table 6). Regardless of feedback type, verifiable public is the most efficient mechanism for principals since the entry rate is almost 100% and it has a significant positive effect on performance ($z= 3.466$, $p= 0.0005$ according to Wilcoxon signed-rank test). In order to rule out period and configuration effect on entry rate, we additionally perform a logit regression of entry on all treatments by controlling period and configuration. The results confirms that in verifiable-public mechanism, the

probability of entry has a strictly increasing tendency while other mechanisms has either negative or no effect on entry rate(see Table 7).

Figure 6: Pre and Post-Feedback Performance in different treatments

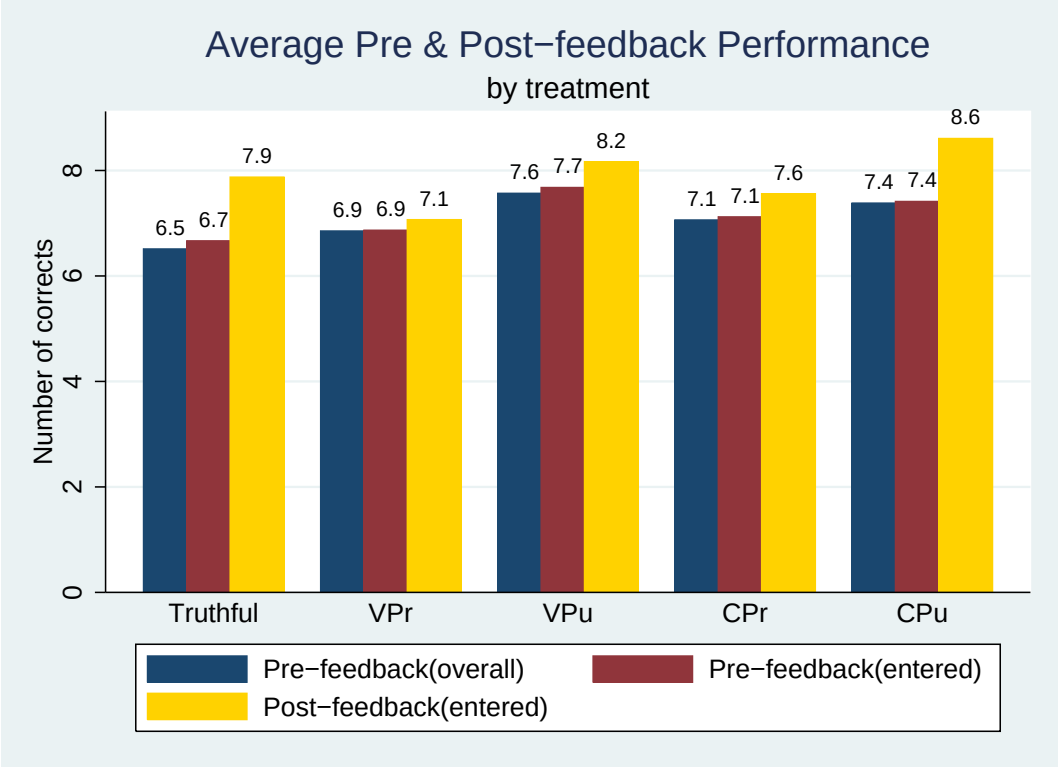


Table 6: Entry rate in different treatments

Treatment	Percentage of entry
Truthful(Baseline)	90.11
Verifiable-Private	76.92
Verifiable-Public	96.70
Cheap-Private	87.91
Cheap-Public	86.75
Overall	87.70

Table 7: Logit regression of entry rate

Logit (robust std. errors)	
Dependent var:	enter
VPr	-1.267***
	(0.462)
VPu	1.147*
	(0.695)
CPr	-0.364
	(0.494)
CPu	-0.286
	(0.497)
Pre-feedback Performance	0.122**
	(0.0589)
Period	-0.275**
	(0.121)
Configuration	YES
N	447
χ^2	23.45
Prob>χ^2	0.0153

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Logit regression with random effects model

We then consider each feedback type to see how agents react to feedback in terms of performance concerns. We exploit GLS regressions of performance change on each feedback type with some additional controls (see Table 8). Subjects tend to perform better after receiving no feedback, while their performance remain unaffected after a positive feedback. This indicates that people are irresponsive in terms of performances to positive feedback. Rather, they improve their perfor-

mance in an ambiguous state. Also, regarding the other factors, optimism of the agent and period does not significantly influence agent's performance.

Table 8: Change in Performance After Feedback

	(1)	(2)
	change	change
Positive Feedback(1,0)	-0.418	-0.538
	(0.380)	(0.389)
No Feedback(1,0)	0.830*	0.765*
	(0.440)	(0.442)
Period	-0.0725	-0.0824
	(0.118)	(0.118)
Session	YES	YES
Δ(Optimism)		0.0141
		(0.0102)
N	310	310
χ^2	27.43	29.42

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

GLS Regressions for Different Feedback Mechanisms

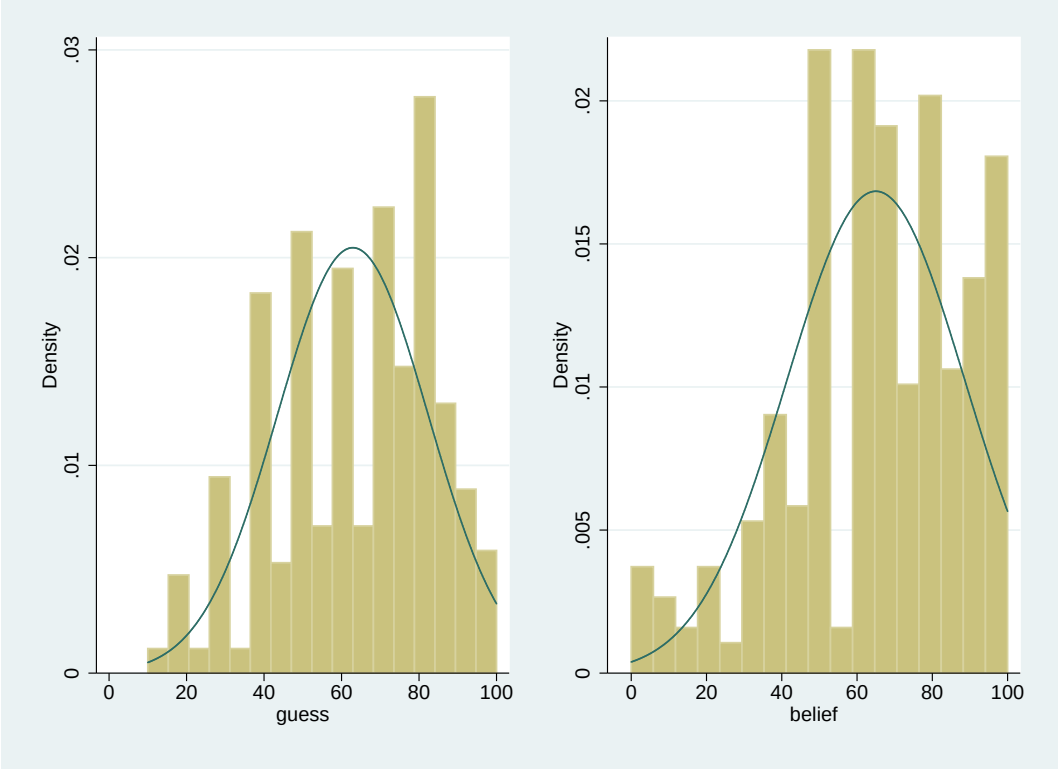
Note that Δ corresponds to change in belief after feedback which implies optimism of subject.

6.2 Principal Side Analysis

Following the agent approach, we now explore the principal side of the analysis. As we mentioned before, investigating the strategy of principal is at the heart of this strand of the study. However, we first examine the expectations of principals from the message processing of the agents in order to understand the rationale behind the strategy. As it is shown in Figure 7, principals expect agents to be less confident(p= 0.08 according to a paired t-test). In order to review this result in a message-specific domain, we test whether the difference between actual posterior

belief of agents and principal's guess are significant for each type of feedback. Here, according to a paired t-test, we find that principals expect positive messages to be interpreted more optimistically than they actually are (p=0.08), whilst they expect negative messages to be evaluated more pessimistically (p= 0.02). Thus, principals generally overestimate the responses of the agents on feedback.

Figure 7: Distributions of principals' guess and agents' post-feedback belief



We then categorize the message portfolio of the principal under different feedback mechanisms. Considering the results in columns VPr and VPu of Table 9, we observe that principals prefer to transmit 44% of the bad outcome in verifiable-private, while they send 56.82% of the bad outcome in verifiable-public feedback mechanisms. A Pearson's chi-square test rejects that the distribution of the message proportions are similar in both contexts ($\chi^2 = 23.8650$, p= 0.000 in VPr and $\chi^2 = 5.7732$, p= 0.016 in VPu). So in both cases, nearly half of the bad outcome is transmitted, while principals are also allowed to withhold the information. There-

fore, there is a general tendency among principals to release the truthful outcome even if it deteriorates their payoff.

Similarly, focusing on bad outcome agents in cheap-talk mechanism on Table 9, principals choose "Successful" message for only 54% of them in private case, whilst they choose that message only for 37.5% of them in public case. The distributions of message frequency is also significantly different in both contexts according to a chi-square test ($\chi^2 = 15.9765$, p= 0.000 in CPr and $\chi^2 = 17.4159$, p= 0.000 in CPu). This indicates that they prefer to send either actual information being "Failed", or no information at all to nearly half of those bad outcome agents. Regarding the stated message proportions in cheap-talk case, the strategy of principal might seem that principals randomize their reports due to the non-credibility concerns. However, since they prefer to send "Successful" message to 94% of the successful agents in Cheap-Private and 85% in Cheap-Public[20], we observe that they actually follow a strategy in those mechanisms, as well. Thereby, we conclude that principals are lie averse since they mostly prefer not to lie even though it affects their own payoff.

Table 9: Message composition of principal under different treatments

	VPr(message)		VPu(message)		CPr(message)			CPu(message)		
Actual	Info	No Info	Info	No Info	S	F	No	S	F	No
S	31	1	31	7	30	0	2	29	2	3
	(96.88)	(3.13)	(81.58)	(18.42)	(93.75)	(0.00)	(6.25)	(85.29)	(5.88)	(8.82)
F	22	28	25	19	27	17	6	15	10	15
	(44.00)	(56.00)	(56.82)	(43.18)	(54.00)	(34.00)	(12.00)	(37.50)	(25.00)	(37.50)

Number of agents received corresponding message is represented in the table

Percentages in parentheses

Here S and F denote "Successful" and "Failed".

In addition to the message composition on independent level, we examine strategy profile in public cases, as well. By the results presented in verifiable-

[20]Each proportion is checked by a test of proportions that it significantly changes according to actual outcome(p= 0.0010 in Cheap-Private and p= 0.0006 in Cheap-Public). This confirms that principals consider the actual outcome in reporting, rather than sending random signals regardless of the outcomes.

public column of Table 10, we observe that there is a significant propensity in principals to reveal truthful information even though the outcome involves two failed agents[21]. The results also indicates that principals always reveal SS outcome which is consistent with our theoretical predictions for verifiable-public case. In cheap-public case, principals prefer to send truthful information to 27% of the agent groups involving one successful and one failed, and to 14% of the groups including both failed agents. A test of proportions rejects that the distributions of reports are same across the actual success outcomes only in SF-FS case (p=0.0560 for SF-FS and p= 0.2485 for FF). Also in FF case, we show that principals choose to transmit no information to 71% of that type agent groups(p= 0 according to a test of proportions). Even though the proportions for both cases are small, there are still at least a portion of principals who prefers to be truthful. In addition, a significant proportion of principals prefers to send no feedback to both agents who failed, where lying is provided as an option as well. Notably, principals significantly refrain from lying.

Table 10: Message composition of principal in public cases

	VPu(message)		CPu(message)					
Actual	Info	No Info	SS	SF-FS	FF	No-No	SNo-NoS	FNo-NoF
SS	16	0	6	0	0	0	2	0
	(100.00)	(0.00)	(75.00)	(0.00)	(0.00)	(0.00)	(25.00)	(0.00)
SF-FS	30	14	22	14	2	0	12	2
	(68.18)	(31.82)	(42.30)	(26.92)	(3.85)	(0.00)	(23.08)	(3.85)
FF	10	12	2	0	2	10	0	0
	(45.45)	(54.55)	(14.29)	(0.00)	(14.29)	(71.42)	(0.00)	(0.00)

Number of agents received corresponding message is represented in the table

Percentages in parentheses

[21]A test of proportions rejects the distribution of the information transmission proportions are same according to actual outcome in both SS and FF cases, while it fails to reject in SF-FS case (p= 0.0571, p= 0.0317, p= 0.9869 for SS, FF and SF-FS cases respectively). Here SF-FS denotes the actual outcome that involves one successful and one failed agents, and SS and FF refers to the outcome which includes both agents are either successful or failed, respectively.

We then additionally check whether principal's strategy affected by other agent's outcome in private case. Note that in practice, the message decision is expected to be independent of success condition of other agent since principals transmit separate feedback to each agent. However, we find a significant negative effect of the success of other agent on the probability of sending any information at all in verifiable-private case(see Table 11, first column). Thereby, although principals are aware of independent message sending process, they seem to be affected psychologically from the other outcome.

Table 11: The effect of actual success of both agents on private message decision

Logit (robust std. errors)	VPr	CPr			
Dependent var:	No Info	Positive Message	Negative Message	Negative Message[1]	No Message
Actual	-3.961***	2.585***	.	-0.334***	-0.656
	(1.080)	(0.797)	.	(0.0688)	(0.873)
Other Actual	-1.159**	0.806	-0.232	-0.0343	-1.556
	(0.585)	(0.578)	(0.629)	(0.846)	(1.104)
Female	0.256	-0.679	0.383	0.0560	0.752
	(0.579)	(0.550)	(0.610)	(0.851)	(0.790)
N	82	82	50	82	82
Psuedo R^2	0.3134	0.2085	0.0099		
R^2				0.1743	
Prob>χ^2	0.0000	0.0001	0.7277		0.1842

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

[1]Note that linear probability model is applied in this regression

Exploring the tendency to releasing bad outcome, we question whether the belief of principal about agent's confidence level bring out that result. We perform several logit regressions of the true message about failure on principal's guess with controlling any gender effect. By the results shown in Table 12, we observe that guess of principal has a significant negative impact on the realization of failure message in all treatments, except Cheap-Public linear probability model. This indicates that as the principal's belief about post-feedback confidence of agent increases, the probability of sending the information about failure drops significantly.

Table 12: Transmission of truthful bad outcome information and principal's belief

Logit (robust std. errors)	Verifiable		Cheap-Talk		
Dependent var: Truebad	Private	Public	Private	Public	Public[1]
Female(1,0)	-0.525	-0.250	0.125	.	0.0462
	(0.817)	(0.969)	(0.974)	.	(0.0429)
Guess	-0.124***	-0.0543***	-0.0864***	-0.0759**	-0.00272
	(0.0294)	(0.0175)	(0.0260)	(0.0312)	(0.00249)
N	82	82	82	32	74
Psuedo R^2	0.39	0.074	0.3172	0.2889	
R^2					0.1259
Prob>χ^2	0.0001	0.0123	0.0039	0.0150	

Entries are logit coefficients. Standard errors are in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Truebad is the dummy that is 1 if truthful bad outcome is sent, 0 otherwise.

[1]Note that linear probability model is applied in this regression

As in the agent case, we supplement a gender-based analysis to our investigation. We first consider lying behavior of principals in gender level and do not find any significant difference at all (p=0.6095 according to a two-sample proportions test). Also from the results exhibited in Table 11 and Table 12, we find no gender effect on our previous results.

7 Conclusion

Interim performance feedback is a widely used motivational tool in business organizations and schools. In application, as well as the truthfulness of the feedback, the audience can vary according to different constraints. Even though it is a prevalent technique in many performance-based environments, the literature is scarce in terms of analysis of both their impact on motivation and on performance in a real-effort setting.

Our study serves as a comparative analysis of four different report mechanisms in which we alter both verifiability of the information, i.e. Verifiable and Cheap-Talk, and the audience of the feedback, i.e. Private and Public. Beside previous studies focused on performance feedback (Ederer and Fehr (2007), Rosaz (2012) and Gürtler and Harbring (2010)), our design provides a novel approach to this strand of the literature from two points. First, we make inference on a real-effort basis which provides more implementable results to real-life settings. Second, we exclude social comparison and identity concerns on performance and belief updating by using an individual-based performance rather than a tournament. However, agents are still influenced by the report about other agent which indicates the existence of psychological factors affecting belief-updating under public reporting.

Our results suggests that the information about the actual outcome is transmitted only in verifiable feedback mechanisms, consistent with our theoretical predictions. Also regarding the efficiency concerns, Verifiable-Public is the optimal mechanism in comparison to the discussed alternatives. By the principal-side results, we observe a general propensity to truth-telling. We additionally present a new insight to literature in terms of understanding the reporting strategy by expectations of the sender. Indeed, the expectations of principals about the agents' self-confidence has a significant effect on reporting strategy, especially on the preference of honesty. We also detect a gender-based difference in interpretation of messages such that women are more optimistic than men when strategically no

feedback is sent. However, the lie aversion of principal is independent of gender since there is no correlation between the reporting behavior and gender.

Our study is focused on a goal-based performance in individual level. Yet, most workplaces are team-oriented such that the target is given to teams rather than individuals. Therefore, to investigate a more applicable real-life setting, we can investigate the same questions in a team-level performance as a further study. We can also implement the same design in a tournament context to check whether our results are generally robust in all performance settings.

References

Abeler, Johannes; Becker, Anke, and Falk, Armin. Truth-telling: A representative assessment. 2012.

Aoyagi, Masaki. Information feedback in a dynamic tournament. *Games and Economic Behavior*, 70(2):242–260, 2010.

Azmat, Ghazala and Iriberri, Nagore. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7):435–452, 2010.

Barankay, Iwan. Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment. Technical report, Working paper, 2011.

Barr, Steve H and Conlon, Edward J. Effects of distribution of feedback in work groups. *Academy of Management Journal*, 37(3):641–655, 1994.

Battaglini, Marco and Makarov, Uliana. Cheap talk with multiple audiences: An experimental analysis. 2010.

Charness, Gary and Dufwenberg, Martin. Promises and partnership. *Econometrica*, 74(6):1579–1601, 2006.

Chen, Ying; Kartik, Navin, and Sobel, Joel. Selecting cheap-talk equilibria. *Econometrica*, 76(1):117–136, 2008.

Crawford, Vincent P. and Sobel, Joel. Strategic information transmission. *Econometrica*, 50(6):pp. 1431–1451, 1982.

Ederer, Florian. Feedback and motivation in dynamic tournaments. *Journal of Economics & Management Strategy*, 19(3):733–769, 2010.

Ederer, Florian P and Fehr, Ernst. Deception and incentives: how dishonesty undermines effort provision. *University of Zurich Institute for Empirical Research in Economics Working Paper*, (341), 2007.

Eil, David and Rao, Justin M. The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138, 2011.

Ertac, Seda. Social comparisons and optimal information revelation: Theory and experiments. *Job Market Paper, UCLA*, 2005.

Ertac, Seda. Does self-relevance affect information processing? experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization*, 80(3):532–545, 2011.

Ertac, Seda and Szentes, Balazs. The effect of performance feedback on gender differences in competitiveness: experimental evidence. *Work. Pap., Koc Univ., Turkey*, 2010.

Farrell, Joseph and Gibbons, Robert. Cheap talk with two audiences. *The American Economic Review*, 79(5):1214–1223, 1989.

Fischbacher, Urs. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178, 2007.

Fischbacher, Urs and Föllmi-Heusi, Franziska. Lies in disguise-an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547, 2013.

Gibbs, Michael J. An economic approach to process in pay and performance appraisals. *University of Chicago, Graduate School of Business*, 1991.

Gill, David and Prowse, Victoria. Gender differences and dynamics in competition: The role of luck. 2010.

Gill, David and Prowse, Victoria. A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review*, 102(1):469–503, 2012.

Gneezy, Uri. Deception: The role of consequences. *The American Economic Review*, 95(1):384–394, 2005.

Grossman, Sanford J. The informational role of warranties and private disclosure about product quality. *Journal of law and economics*, 24(3):461–483, 1981.

Gürtler, Oliver and Harbring, Christine. Feedback in tournaments under commitment problems: Experimental evidence. *Journal of Economics & Management Strategy*, 19(3):771–810, 2010.

Kartik, Navin. Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395, 2009.

Kartik, Navin; Ottaviani, Marco, and Squintani, Francesco. Credulity, lies, and costly talk. *Journal of Economic theory*, 134(1):93–116, 2007.

Koessler, Frédéric. Lobbying with two audiences: Public vs private certification. *Mathematical Social Sciences*, 55(3):305–314, 2008.

Milgrom, Paul R. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391, 1981.

Mobius, Markus M; Niederle, Muriel; Niehaus, Paul, and Rosenblat, Tanya S. Managing self-confidence: Theory and experimental evidence. Technical report, National Bureau of Economic Research, 2011.

Murphy, Kevin R and Cleveland, Jeanette N. *Performance appraisal: An organizational perspective.* Allyn & Bacon, 1991.

Rosaz, Julie. Biased information and effort. *Economic Inquiry*, 50(2):484–501, 2012.

Sanchez-Pages, Santiago and Vorsatz, Marc. Enjoy the silence: an experiment on truth-telling. *Experimental Economics*, 12(2):220–241, 2009.

Seidmann, Daniel J and Winter, Eyal. Strategic information transmission with verifiable messages. *Econometrica: Journal of the Econometric Society*, pages 163–169, 1997.

A Appendix

A.1 Appendix 1. Instructions

[Original Instructions were in Turkish and they are available upon request.]

Welcome. Thank you for participating in our experiment. The experiment you are going to take part is about economic decision making. During the study, you are expected to solve some questions in a test and make decisions regarding your performance on that test. You will earn 10 TL for your participation. Besides the show up-fee, your earnings in the experiment will depend on your performance, accuracy of your decisions and chance. There is not any kind of misleading or deception in this study. The rules will be explained soon are totally correct and your payment will be determined according to them.

Participation in this study is entirely voluntary. After the rules explained, you are free to quit until the experiment starts. Any of your decisions throughout the experiment will not be matched with your real identity, but rather will be recorded under an anonymous subject number.

Experiment consists of 5 periods and within each period, there are 2 rounds. At the end of the experiment, one round out of 10 rounds in total will randomly be selected and your earning will be according to your payoff in that chosen round. Payments will be made privately in cash at the end of the experiment.

During the experiment, all monetary activities will be done in ECU ("experimental currency unit") where 1 ECU equals to 0.06 TL.

For each of the 5 periods, different rules will be applied. Just before each period, the rules of the corresponding period will be explained in detail.

Primarily, the decision activity which you will encounter in each rounds will be explained. After you understand this part, the instructions about each of the 5 periods will be explained period by period.

General Rules

At the beginning of the experiment, participants will be appointed randomly

to one of the roles of "Principal" and "Agent". The roles will be fixed throughout the experiment. Now, we are going to explain the rules that will be applied for agents.

Agents:

Agents will perform a task in each period. There are 2 types of tasks (which will be appointed randomly):

i. Addition task

ii. Verbal task

In addition task, agents are expected to solve the addition questions consisting 4 or 5 two-digit numbers in 120 sec. (e.g. 11+ 48+ 96+24=?)

In verbal task, agents are expected to solve in 120 sec following type of questions.

- General knowledge questions(Geography, Literature, Sports...)
- Verbal classification questions
- Number-letter matching questions

The questions for each type in verbal task is as following:

- General knowledge: questions from diverse fields. For example:
  - What is the capital city of Canada?
  - Who is the writer of the novel *Oliver Twist*?
- Verbal classification questions: the questions are based on grouping according to the relationship between them. For example:

  - Which of the following choices include the relationship between $Tea : Sugar$?
  - When you classified the following words, which one is odd?

- Number-Letter matching questions: we assign each letter to a number and ask in the following way:

  - ALI= 123
  - ILE= 324
  - AILE=? where the answer is AILE= 1324.

For each period, a "target score" will be determined for each task separately. The outcome will be determined as, if the number of correctly solved questions is:

- higher than or equal to the target score, *Successful*,

- lower than the target score, *Failed*

The payoffs of being Successful of Failed will be as following:

If you are Successful (you pass the target score), you will earn 300 ECU. If you are Failed (cannot pass the target score), you will earn 100 ECU.

The target score will be keep fixed through both rounds of the corresponding period. Your number of correctly solved questions in both rounds within a period will be compared with the same target score. However, you will not be informed about the target score.

The task you will perform will be the same in the both rounds within a period. However, the difficulty of the questions might vary between rounds.

After first round, for the second round, agents can choose either to determine their second round payoff by their second round performance or to leave it to the chance mechanism. As an agent:

- If you choose performance mechanism for second round,
  - If Successful in second round performance (pass the target score) 300 ECU,
  - If Failed in second round performance (cannot pass the target score) 100ECU.
- If you choose chance mechanism for second round,
  - You will not perform in that round.
  - You will earn, 300 ECU with probability X, and 100 ECU with probability (100-X).

You will make this performance/chance mechanism decision by answering the following question:

What is the minimum probability of chance (X) to leave your second-round payoff to chance mechanism rather than your own performance? (%)

In other words, you are expected to state in which probabilities you choose your own performance and in which probabilities you prefer chance mechanism. With this decision, you will also state indirectly that how you are confident about your success in second period.

The probability of winning for chance mechanism (X) will be determined randomly by computer. If you choose performance mechanism at the chosen chance

probability, then you will perform in second round. Otherwise, rather than your performance, your payoff will depend on chance (You will earn 300 ECU with probability X% and 100 ECU with probability (100-X)%).

Now, we have a short quiz about the rules we have explained. Your answers will not affect your earnings. Rather, they are aimed to check your understanding of the rules. Please do not hesitate to ask questions if you do not understand a particular detail. Also keep in mind that any kind of communication with other participants is prohibited.

[Quiz-see Appendix 2]

Agents will give the performance/chance mechanism decision twice: before and after they received the feedback their performance outcome (Successful or Failed). One of those stated minimum probabilities will randomly be chosen and applied.

The time-line of the agents will be as following*[shown from slides]*

1. (1^{st} round) The appointed task will be performed

2. The decision about performance/chance mechanism for 2^{nd} round will be stated

3. Feedback about the 1^{st} round will be given (Successful/Failed)

4. The pre-feedback performance/chance decision will allowed to be changed (if you want you can state the same amount)

5. One of the pre- and post-feedback decisions will randomly be chosen. That decision will be compared the chance probability X and performance or chance mechanism will be applied in 2^{nd} round accordingly.

6. (2^{nd} round) If performance mechanism chosen: As in the first round, the same type of task will be performed and your earnings will depend on your success condition (300 ECU/100 ECU).

If chance mechanism chosen: No performance. 300 ECU with probability X and 100 ECU with probability (100-X) will be earned.

Principals:

Each period, 3-person groups which will be consisted of 1 Principal and 2 Agents will be formed. At the beginning of each period, groups will be shuffled and formed repeatedly. The agents in the same group will always perform different type of tasks and will be evaluated according to different target scores.

Principal type participants will not perform any task. However, they can see the questions solved by both agents on their screen.

The payoffs of principals(valid for all periods):

1^{st} round of each period: 100 ECU constant payoff

2^{nd} round of each period: payoff will be dependent upon agents' decisions and performances.

Return of one agent to principal in second round($return_i$):

- If the agent chose performance mechanism in second round:

  $return_i$= Number of corrects in second round * 20 ECU

- If the agent chose chance mechanism: $return_i$= 0 ECU

Payoff function of principal for second round= *50ECU(constant payoff)*+$min(return_1, return_2)$,

where $return_1$ and $return_2$ refers to returns from the agent 1 and agent 2.

Therefore if at least one of the agents chooses chance mechanism, principal will not earn a payoff from agents; so for gaining a payoff from agents' performance, both of the agents must perform in the second round.

In addition to this, principals are expected to guess the stated minimum probabilities(%). Principals can see the success outcomes of agents and guess the post-feedback performance decisions of both agents separately. If your guess will be in the ± 5 interval of agent's actual decision, you will earn 10 ECU bonus for each correct guess. At the end of the experiment, your extra earning in the chosen round will be added to your payoff from that round.

Do you have any question?

Part I:

As we explained before, in this part 3-person groups, 1 Principal 2 Agents, will be formed.

Agents will observe their success outcome (Successful/Failed) before they make post-feedback performance/chance decision. This feedback will be sent by computer and it will be certainly true information.

In this part, principals will only see the questions that agents are solving and observe the success outcomes of the agents. Principals are only expected to guess the agents' performance/chance decision in this part.

Do you have any question?

Part II:

In this part, agents will receive feedback about the success outcome from principals. The information from principal must be true. Though, principals also have the option to hold the information.

The message from principal will be privately sent. In other words, principals can choose different messages for agents and agents will only see the message about self-success.

[Screen-shots of both principal and agent screens were shown in slides. They are available upon request.]

Do you have any question?

Part III:

In this part, agents will receive feedback about the success outcome from principals. The information from principal must be true. Though, principals also have the option to withhold the information.

The message from principal will be publicly sent. Principal will choose either sent truthful information to both of the agents or send no information to none of them.

If principal chooses to send the info, agents will see the outcome of other agent along with their own success outcome. Note that the other agent performs another type of task and evaluated with a different target score.

Do you have any question?

Part IV:

In this part, agents will receive feedback about the success outcome from principals. The information from principal does not need to be true. Yet, principals also have the option to hold the information.

Principals will observe the actual success outcomes of agents. Following that, they will choose to send a message among "Successful", "Failed" and "No Info" options regardless of actual outcomes.

The message from principal will be privately sent. In other words, principals can choose different messages for agents and agents will only see the message about self-success.

Do you have any question?

Part V:

In this part, agents will receive feedback about the success outcome from principals. The information from principal does not need to be true. Yet, principals also have the option to hold the information.

Principals will observe the actual success outcomes of agents. Following that, they will choose to send a message among "Successful", "Failed" and "No Info" options regardless of actual outcomes.

The message from principal will be publicly sent. Agents will see the outcome of other agent along with their own success outcome. Note that the other agent performs another type of task and evaluated with a different target score.

Do you have any question?

The decisions to be made are over. Now you are kindly asked to answer survey questions that will be on your screen soon. Thank you again for your participation.

A.2 Appendix 2. Quiz

1. Assume that you state 60% as the probability of being successful. Computer picked the probability of chance mechanism as 55%. Which one of the following is true?

   (a) My probability of earning 300 ECU is 55% and my earning is independent of my performance.

   (b) My earning is 300 ECU if I will be successful and 100 ECU otherwise.

2. Assume that you state 30% as the probability of being successful. Computer picked the probability of chance mechanism as 55%. Which one of the following is true?

   (a) My probability of earning 300 ECU is 55% and my earning is independent of my performance.

   (b) My earning is 300 ECU if I will be successful and 100 ECU otherwise.

3. Person A stated 80% as the minimum probability of leaving her payoff while Person B stated 45% for the same decision. Which one of those people is more self-confident about their success?

   (a) A

(b) B

A.3 Appendix 3. Post-experiment Survey Questions

1. How old are you?
2. What is your gender?[22]
3. What is your year/class?[23]
4. Which one is your faculty?[24]
5. What is your current GPA?[25]
6. Was the experiment and the rules understandable? Please answer on a scale of 1 to 10: 1= not understandable at all, 10= absolutely understandable.
7. How difficult was the addition task? Please answer on a scale of 1 to 5: 1= not difficult at all, 5= extremely difficult.
8. How difficult was the verbal task? Please answer on a scale of 1 to 5: 1= not difficult at all, 5= extremely difficult.
9. As agents, the decision for performance/chance mechanism, is the thought of " if I wait, I will get bored" affected your decision?(Yes/No)
10. Aside from monetary payoffs, how important was it for you to be "Successful"(passing the target score)? [26]

[22]Female participants are denoted as 1, whereas male participants are denoted as 0.

[23]Seniority variable which is recorded as following: 0 for prep, 1-4 for corresponding year(Freshman-Senior), 5 for master's and 6 for PhD students.

[24]1 is for Engineering, 2 stands for CASE, 3 corresponds to Science, 4 is for Law, 5 is for Art and Design whereas 6 stands for Medicine/Nursery.

[25]GPA is recorded as 6 categorical intervals: 1 for below 2.00, 2 for 2.00-2.49, 3 for 2.50- 2.99, 4 for 3.00-3.49, 5 for 3.50- 4.00 and 6 for N/A.

[26]1 for Not important at all, 2 for Not important, 3 for Neither important, nor unimportant, 4 for Important, 5 for Extremely important.