Spring 4-28-2020

Silent or Salient? Ability Heterogeneity in Tournaments

HAO HE
GSU

Silent or Salient? Ability Heterogeneity in Tournaments

BY

Hao He

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Doctor of Philosophy

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2020

ACCEPTANCE

This dissertation was prepared under the direction of the *Hao He* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. Michael Majerczyk, Co-Chair
Dr. Ivo Tafkov, Co-Chair
Dr. Curtis Mullis, Member
Dr. Tyler Thomas, External Member (University of Wisconsin Madison)

ABSTRACT

Silent or Salient? Ability Heterogeneity in Tournaments

BY

Hao He

April 29, 2020

Committee Chair: *Dr. Michael Majerczyk (Co-Chair), Dr. Ivo Tafkov (Co-Chair)*

Major Academic Unit: *School of Accountancy*

I experimentally investigate the impact of the level and salience of the ability heterogeneity on the effectiveness of relative performance information (RPI) on individual performance in tournaments. In my setting, the first stage is used to sort out the task abilities while the second stage consists of a tournament. My primary dependent variable is the effectiveness of the RPI provided at the end of the first stage, captured by the change of individual performance over the stages. Consistent with my predictions, I find that RPI's effectiveness is more positive when the ability heterogeneity is low and salient than when it is high and salient, because the knowledge of ability heterogeneity influences participants' performance expectations. I also find that RPI's effectiveness is more positive when the ability heterogeneity is low and salient than when it is not salient, because both high and low performers attach greater value to the tournament outcomes in the former situation. Moreover, I find that RPI's effectiveness is not more negative when the ability heterogeneity is high and salient than when it is not salient, as unequal competitions might give rise to additional value in winning the tournament. Finally, I find that performance reduction in low performers is smaller than that in high performers, once the ability heterogeneity become salient and low. My results suggest that when RPI is present, its effectiveness depends on the ability heterogeneity and its saliency that firms should consider when designing effective tournaments.

Silent or Salient? Ability Heterogeneity in Tournaments

Hao (Alan) He

Georgia State University

April 2020

I am indebted to my dissertation committee co-chairs, Michael Majerczyk and Ivo Tafkov, for their continual help and guidance. I thank members of my dissertation committee – Curtis Mullis and Tyler Thomas for their invaluable suggestions and insights. I am grateful for my wife Ruyun Feng, for her support. I also thank Milton Shen, Kirill Novoselov, and workshop participants at Georgia State University, Arkansas State University, University of Alabama-Huntsville, University of Northern Georgia, University of Southern Maine, Eastern Connecticut State University, and Central Connecticut State University for helpful comments. Any errors are my own.

Silent or Salient? Ability Heterogeneity in Tournaments

Abstract

I experimentally investigate the impact of the level and salience of the ability heterogeneity on the effectiveness of relative performance information (RPI) on individual performance in tournaments. In my setting, the first stage is used to sort out the task abilities while the second stage consists of a tournament. My primary dependent variable is the effectiveness of the RPI provided at the end of the first stage, captured by the change of individual performance over the stages. Consistent with my predictions, I find that RPI's effectiveness is more positive when the ability heterogeneity is low and salient than when it is high and salient, because the knowledge of ability heterogeneity influences participants' performance expectations. I also find that RPI's effectiveness is more positive when the ability heterogeneity is low and salient than when it is not salient, because both high and low performers attach greater value to the tournament outcomes in the former situation. Moreover, I find that RPI's effectiveness is not more negative when the ability heterogeneity is high and salient than when it is not salient, as unequal competitions might give rise to additional value in winning the tournament. Finally, I find that performance reduction in low performers is smaller than that in high performers, once the ability heterogeneity become salient and low. My results suggest that when RPI is present, its effectiveness depends on the ability heterogeneity and its saliency that firms should consider when designing effective tournaments.

Keywords: heterogeneity; salience of information; relative performance information; tournament

1. Introduction

This study investigates how the level and salience of employee heterogeneity affects the effectiveness of relative performance information (RPI) in tournaments. Firms commonly use tournaments to motivate employees (Berger, Libby, and Webb 2018; Chen, Williamson, and Zhou 2012; Hazels and Sasse 2008; Kwoh 2012). In tournament settings, employees often receive feedback about their relative performance (Nordstrom, Lorenzi, and Hall 1990; Gino and Staats 2011; Gill, Kissova, Lee, and Prowse 2018). As accountants play a major role in providing information and feedback to employees (Bonner and Sprinkle 2002; Hannan, McPhee, Newman, and Tafkov 2013), it is important for accountants to understand conditions where the effectiveness of RPI (i.e., positive effect of RPI on task performance) varies.

Firms such as General Electric, Yahoo!, and Whirlpool are continuously experimenting with different forms of relative-performance feedback (Kuhnen and Tymula 2012). This experimenting may be the result of a lack of consensus about the effectiveness of RPI in the workplace (e.g., Prendergast, 1999; Grote 2005; Hazels and Sasse 2008). The extant literature has yielded mixed evidence on the effectiveness of RPI in tournaments. For example, Eriksson, Poulsen, and Villeval (2009) find that RPI induces higher performance in all employees. Kuhnen and Tymula (2012), however, provide evidence that top performers shirk with RPI and Fershtsman and Gneezy (2011) report lower performance as quitting becomes more common with RPI. Delfgaauw, Dur, Non, and Verbeke (2014) find no overall effect of RPI as the high performers improve performance while the low performers do not respond to tournament incentives. The inconclusive findings might be related to performance levels which vary across a continuum (Hamilton, Nickerson, and Owan 2003). More so, the mixed evidence might be attributable to employees' perceptions of ability heterogeneity, which could depend on the feedback that makes

the perceptions more salient. To the extent that such perceptions are important to the responses to feedback (Abeler, Falk, Goette, and Huffman 2011; Gill et al. 2018; Perez-Truglia 2019), the impact of RPI on performance might be contingent on whether the knowledge of ability heterogeneity is salient and the level of ability heterogeneity.

In management accounting, Hannan, Krishnan, and Newman (2008) find that RPI may weaken mean performance under tournament contracts. Hannan et al. argue that employees who are lagging in performance learn from a series of precise RPI that winning is unlikely and as a result work less efficiently. However, depending on the level, and salience of employee heterogeneity, RPI may influence employees differentially. Specifically, when the actual level of employee heterogeneity is uncertain, the high performers might attribute their relative standing to perceived higher abilities while the low performers might attribute their relative standing to a belief that they put forth lower effort. As they receive more updates on RPI, perceived distribution of task ability might become salient and heterogeneous (Casas-Arce and Martinez-Jerez 2009; Ederer 2010). When the level of employee heterogeneity is high and salient, both high and low performers are more likely to attribute their relative standing to the ability advantage or disadvantage, resulting in effort reduction. When the level of employee heterogeneity is low and salient, performance is more likely to be attributed to high or low effort level, resulting in equal or even higher effort and, thus, performance.

In this paper's setting, participants' abilities are sorted out before they compete in a tournament that determines the winner. The setting is generalizable to the real world as firms often sort employees based on their abilities during hiring and promotion processes (Lise, Meghir, and Robin 2016; Milgrom and Roberts 1992). When the ability heterogeneity is salient, I argue that a low level of ability heterogeneity likely homogenizes tournament participants' performance

expectations and might augment the non-pecuniary value of rank by intensifying social comparison (Festinger 1954; Kulik and Ambrose 1992). Hence, I predict that when RPI is present task performance is highest when the ability heterogeneity is low and salient but is lowest when it is high and salient.

When the ability heterogeneity is not salient, I argue that participants are more uncertain about their tournament outcomes compared to when the ability heterogeneity is high and salient, which encourages exerting more effort. Participants are also less likely in both of these conditions to associate the same value to rank compared to when the ability heterogeneity is low and salient. As such, I predict that when RPI is present, task performance is higher when the ability heterogeneity is not salient compared to when it is high and salient, but lower compared to when it is low and salient. In addition, I argue that while both high and low performers likely will increase performance when the ability heterogeneity is low and salient, the self-serving attributions will likely lead to a differential impact on performance between the two types of participants. Therefore, I predict that when RPI is present, the drop in the performance change from when the ability heterogeneity is low and salient to when it is not salient, is higher for high performers than low performers.

To test my predictions, I collect data using a tournament setting in which the first stage of the tournament is used to sort out the task abilities while the second stage determines the winners. I adapt the task design from Chan (2018) which is an advanced slider task with both an ability part that requires solving subtraction problems and an effort part that requires positioning sliders. I manipulate the level and salience of ability heterogeneity using a between-participants design with three conditions: *Homogeneous* versus *Heterogeneous* versus *NotSalient*. In *Homogeneous* and *Heterogeneous* conditions, participants are made aware that they compete in the second stage with

peers of similar or different task ability based on RPI presented at the end of the first stage. In *NotSalient* condition, participants are not informed about the task ability of the peers they compete in the second stage relative to their own. I examine participants' behavior before and after they receive the RPI across the three conditions. My primary dependent variable, is the difference of individual performance between the second and the first stage scaled by the first-stage performance, which captures the effectiveness of RPI.

I find that RPI's effectiveness is more positive in *Homogeneous* than *Heterogeneous* condition, and the effect is driven by the knowledge of ability heterogeneity influencing participants' performance expectations and the achievement goal of the participants. Consistent with my predictions, I also find that RPI's effectiveness is more positive in *Homogeneous* than *NotSalient* condition. Additional analysis confirms that low performers attached greater non-monetary value to tournament outcomes in *Homogeneous* condition, and that perceived heterogeneity drives the difference in RPI's effectiveness between the conditions. Moreover, I find that RPI's effectiveness is not more negative in *Heterogeneous* than *NotSalient* condition. Additional analysis confirms that while low performers in *Heterogeneous* condition are demotivated due to low winning expectations, they're also motivated by greater non-monetary value of ranks. Finally, I find that the drop in RPI's effectiveness from *Homogeneous* to *NotSalient* condition is higher for high performers than low performers. The effect is driven by high performers in *NotSalient* condition self-servingly attributing their relative performance to the ability advantage while low performers equally attribute their relative performance to ability disadvantage and insufficient effort.

Results of this study underscore the inter-related nature of tournament structure and information system design choices and have important implications for the design of effective

tournaments. These results contribute to a stream of research that has only recently begun to consider the role that tournament structure plays in determining the effect of RPI on performance in tournaments (Freeman and Gelber 2010; Newman and Tafkov 2014). Particularly, the study extends prior research (Newman and Tafkov 2014) that shows that the negative effect of providing RPI in tournaments, as documented by Hannan et al (2008), is not universal to all tournament structures. By showing that that the effectiveness of providing RPI varies based on the level and salience of tournament participants' ability heterogeneity, I highlight that firms should keep in mind the structure of their tournaments when providing RPI. Specifically, my results suggest that when RPI is present firms should provide information about the tournament structure in terms of the level of ability heterogeneity from employees under a homogeneous tournament but may withhold such knowledge under a heterogeneous tournament. Also, the results suggest that if RPI is readily available via informal sources (Hannan et al. 2013), the effectiveness of tournaments may be affected by perceived ability heterogeneity among the participants. However, it is not necessarily the case that firms need to discontinue the tournaments to avoid negative performance effects if the informal information suggests more homogeneous tournaments.

The remainder of this paper proceeds as follows. Section 2 provides background information and develops my hypotheses; Section 3 describes the experiment; Section 4 reports the results; and Section 5 concludes the study.

2. Theory and Hypothesis Development

2.1 Background

How individuals respond to RPI in tournaments is empirically inconclusive (Kluger and Denisi 1996; Smither, London, and Reilly 2005). Even though it can encourage some to catch up

or to excel, others may be demotivated or become complacent. For example, in a lab experiment, Eriksson et al. (2009) find that RPI induces higher performance in both high and low performers. The perceptions of ability heterogeneity among the participants was not measured or manipulated in their study, which might have contributed to the findings. In contrast, Fershtsman and Gneezy (2011) report that RPI is associated with high performers reducing effort and a substantial portion of low performers giving up. Although ability heterogeneity was manipulated in their field experiment (similar to *Homogeneous* condition), the experiment was conducted with teenagers competing in 60-meter races. The observed quitting behavior might be attributable to the subjects not taking running ability seriously. Other studies find that high and low performers respond to RPI in opposite directions. For instance, Hannan et al. (2008) find that high performers increase performance while low performers work inefficiently, resulting in lower overall tournament performance after receiving RPI. Similarly, in a field experiment involving several geographically scattered sales teams, Delfgaauw et al. (2014) find no overall effect of RPI as the high performers improve performance while the low performers do not respond to tournament incentives.

Relatedly, in non-tournament settings, the impact of RPI is also not as clear. A number of papers find higher performance or greater effort with RPI (Gneezy and Rustichini 2004; Mas and Moretti 2009; Freeman and Gelber 2010; Tafkov 2013; Azmat and Iriberri 2016; Bradler, Neckermann, and Non 2016). However, some papers report lower performance (Bellemare et al. 2010; Barankay 2011, 2012; Bandiera et al. 2013; Ashraf, Bandiera, and Lee 2014) or find no clear pattern (Gino and Staats 2011; Bhattacharya and Dugar 2012; Rosaz, Slonim, and Villeval 2012; Georganas, Tonin, and Vlassopoulos 2015). Most of these studies do not explicitly consider ability heterogeneity as a possible factor in influencing how individuals respond to RPI. To this point in the literature, the effects of ability heterogeneity on RPI's effectiveness are largely unexplored.

Prior literature also provides some evidence suggesting high performers in tournaments might be susceptible to complacency effect (Cardinaels, Chen, and Yin 2018). Cardinaels et al. (2018) and Casas-Arce and Martinez-Jerez (2009) find that contestants tend to be complacent and reduce their subsequent effort when RPI indicates they are in a leading position. Berger et al. (2013) analyze repeated tournaments and find immediate complacency in the winners after the prize is awarded. In a real corporate setting, Cai and Gallani (2018) also observe a similar complacency effect in high performers. However, Muller and Schotter (2009) find that high performers are motivated to work harder while learning their leading position. In high-skill settings where participants are of high dispositional ability, non-complacency is more prevalent, but the evidence is also mixed. Using data from NBA games, Berger and Pope (2011) find that leading in the first half is associated with reduced effort. In professional golf tournaments, however, Guryan, Kroft, and Notowidigdo (2009) find no effect of RPI on effort or performance. In addition to providing mixed evidence of complacency in high performers, whether dispositional or situational, the aforementioned studies also do not take into account the effects of ability heterogeneity.

2.2 Hypotheses

Ability heterogeneity refers to the difference in the task-related ability among tournament participants. Within a tournament, task-related ability naturally varies across a continuum (Hamilton et al. 2003), ranging from extreme homogeneity (or very low levels of ability heterogeneity) to an extreme level of ability heterogeneity. Given ability heterogeneity will likely shape tournament participants' performance expectations (Abeler et al. 2011; Gill et al. 2018; Perez-Truglia 2019), I predict the effectiveness of RPI will be contingent on the knowledge of ability heterogeneity.

Specifically, when the heterogeneity level is high and salient, winning is perceived ex ante as more likely by the high performers and less likely by the low performers. This is because, ceteris paribus, the level of task ability is positively associated with task performance, as studies in both economics and psychology find that task abilities are stable and predict performance (Heckman, Jagelka, and Kautz 2019). Employees are often effort-averse (Akerlof and Yellen 1990; Lambert 2007). As a result, high performers will likely calibrate their effort level based on RPI such that they can win the tournament without wasting costly effort. Realizing the ability disadvantage, low performers perceive the tournament as an uphill battle and are thus inclined to save effort. Therefore, making salient the high level of ability heterogeneity is likely to demotivate both high and low performers from exerting effort, resulting in low performance. When the heterogeneity level is low and salient, however, both winning and losing are perceived as equally likely by all participants. Within each tournament group, then, they do not perceive themselves ex ante as either high or low performers. As any differences in individual ability no longer drive the differences in relative performance here, participants must rely on exerting effort to improve their ranks. Within each tournament group, all the participants might perceive their positions under constant threat from each other, and thus are more likely to maintain or increase than to decrease effort.

RPI influences tournament participants through two channels: 1) marginal pecuniary return of effort, and 2) non-pecuniary value of rank (Gill et al. 2018). RPI facilitates social comparison among the participants as it allows individuals to compare themselves to others (Tafkov 2013). Prior accounting literature provides evidence that the motivational effect of social comparison exists even when compensation is not tied to peer performance (e.g., Hannan et al. 2008). When ability heterogeneity is low compared to when it is high, individuals are more likely to engage in social comparison since the comparison target (i.e., the other person) is perceived as more similar

on task ability (Garcia and Tor 2007). The greater engagement in social comparison enhances individual motivation to outperform others (Brown, Ferris, Heller, and Keeping 2007; Brown and Gallagher 1992). Therefore, making salient that the competition is among peers of similar ability can cause participants to attach greater non-pecuniary value to winning. Specifically, in a tournament full of high performers, the participants attach greater nonmonetary value to winning since winning among the top performers suggests a higher status than winning among performers with a range of skill (Altmann, Falk, and Wibral 2012; Dohmen, Falk, Fliessbach, Sunde, and Weber 2011; Moldovanu, Sela, and Shi 2007). In a tournament full of low performers, losing brings additional nonmonetary loss of being identified as the worst in the organization. Prior research suggests that being among the worst increases the participant's need for proving otherwise (Gill et al. 2018). Taken together, making salient the low level of ability heterogeneity might encourage participants to increase effort, likely resulting in performance improvement. In sum, the effectiveness of RPI is higher when participants are aware that they are competing with peers of similar level of ability than with peers of different levels of ability.

H1: When ability heterogeneity is salient, RPI's effectiveness is higher when the level of ability heterogeneity is low compared to when it is high.

The effectiveness of RPI might also be affected by the salience of ability heterogeneity. Individuals are poor information aggregators. They do not consider fully all the information available to them, but rather over-emphasize the information on which their minds focus (Bordalo, Gennaioli, and Shleifer 2012; Gennaioli and Shleifer 2010). In psychology, salience detection is a key mechanism enabling humans to focus their limited cognitive resources on a relevant subset of the available information (Taylor and Thompson 1982).

Firms often have an informational advantage about the ability heterogeneity of participants in the tournament, especially when the performance difference is not directly observable, or when the performance appraisal involves multiple skills or has some subjective element (Ederer and Fehr 2017). Firms sometimes form and make known more homogeneous tournaments, in which case, as previously discussed, RPI causes a stronger preference for rank. First, both internal and external hiring and promotion processes are considered rank-order tournaments (Chan 1996). When hiring or promoting talents from inside a company, the candidates have at least some knowledge of each other's ability (Waldman 1984; Greenwald 1986), ability heterogeneity is low and salient. The inclusion of external candidates, in comparison, makes the heterogeneity less salient and the tournaments intrinsically more heterogeneous (Bidwell 2011). Second, companies sometimes segment a tournament by task ability. For example, tournament platforms such as TopCoder divide the competitors into two competitor pools according to an ability level cutoff, and make both individual ability and performance information available (Boudreau, Lakhani, and Menietti 2016; Heite 2020). More generally, labor forces in reality are sorted into corporate tournaments of different stakes based on ability (Lazear and Rosen 1981, fn. 5). By making salient the level of ability heterogeneity and participant's own ability the participants tend to react toward such information.

Before a tournament starts, individuals likely form expectations of their chances of winning based on the perceived standing of their ability relative to other participants. When the distribution of ability in the tournaments is known, the expectations serve as a clear reference point in determining their effort provision (Koszegi and Rabin 2006; Abler et al. 2011).[1] When the ability

[1] In homogeneous (heterogeneous) tournaments, both high and low performers should expect that spending effort is (not) worthwhile, compared to when the ability heterogeneity is not disclosed.

heterogeneity is not salient, participants likely will discount the informativeness of such prior perceptions of their ability for the subsequent contest as they cannot infer with certainty their relative standing in the distribution of ability. In other words, compared to when the ability heterogeneity is salient, participants are less likely to perceive that a stable "hierarchy of performance" is formed (Kuhnen and Tymula 2012). Hence, all participants likely will be more uncertain about the subsequent tournament's outcome than when the level of ability heterogeneity is high, but less uncertain than when the ability heterogeneity is low. Particularly, compared to when the ability heterogeneity is low and salient all participants are less likely to feel the threat from peers and attach less value to the tournament outcome. Compared to when the ability heterogeneity is high and salient, however, to decrease the uncertainty of their tournament outcomes, high performers are more likely to exert effort to preserve the current perceived position while low performers are more likely to exert effort in an attempt to improve their standing as winning remains possible. In sum, it is expected that when the ability heterogeneity is not salient, RPI's effectiveness is higher than when the heterogeneity is high and salient but lower than when it is low and salient.

Besides the pecuniary effects discussed above, ability heterogeneity might simultaneously affect the motivation of the participants through non-pecuniary channels. Compared to when ability heterogeneity is not salient, tournaments may arouse additional emotions (e.g., joy, pride, anger, etc.) when it is high and salient (Krakel 2008; Parco, Rapoport, and Amaldoss 2005; Sheremeta 2010). Also, as social comparison is stronger when relative ability is made salient, participants with a lower relative standing might increase effort to catch up in their ranking which indicates status (Piazza and Castellucci 2014). Prior literature suggests that tournaments in organizations often provide status competition and social recognition opportunities, as status is

often pursued as an end in itself and whose effect on monetary incentive is possibly substitutive or complementary (Moldovanu et al. 2007; Piazza and Castellucci 2014; Frey 2007). Compared to when ability heterogeneity is low and salient, the motivation based on emotions is lower when the heterogeneity is high and salient. Prior literature suggests that while emotions motivate participants when the ability difference is low, they both motivate and demotivate participants when the difference is high but with a net motivating effect (Krakel 2008).[2] Therefore, when the heterogeneity is high and salient compared to when it is not salient, participants might be encouraged to increase effort due to evoked emotions and status-seeking intentions, while simultaneously being discouraged by a low winning likelihood. Overall, explicitly pitting participants against unequal peers might not be more discouraging than withholding information of ability heterogeneity.

H2a: RPI's effectiveness is higher when the ability heterogeneity is low and salient than when it is not salient.

H2b (null): RPI's effectiveness is not different when the ability heterogeneity is not salient compared to when it is high and salient.

Even if participants rely on RPI to make effort choices for the subsequent contest, when the ability heterogeneity is not salient, they might self-servingly interpret the heterogeneity level based on past performance. As discussed above, I theorize that participants of different ability will attribute their relative standing to different factors. In particular, high performers will likely attribute their high ranks in the past to high ability. Therefore, compared to when the ability

[2] The motivational effect provided by social comparison between *Homogeneous* and *Heterogeneous* condition is discussed in the theory development of H1. Taken together, emotions and social comparison motivate participants most in *Homogeneous* condition, and least in *NotSalient* condition.

heterogeneity is low and salient, high performers are more likely to rely on a perceived advantage in ability to save costly effort.

Similarly, low performers will likely attribute their low ranks in the past to low effort, to maintain a positive self-image (Lazarus 1991; Smith 2000). In addition, due to a persistent bias of overconfidence in relative ranking (Moore and Healy 2008; Alicke, Vredenburg, Hiatt, and Govorun 2001), low performers likely will maintain the expectation of winning in the subsequent tournament insofar as the ability heterogeneity remains non-salient. Therefore, compared to high performers, low performers are less likely to refrain from exerting effort as a persistent winning expectation might drive them to do so. Thus, I predict that the low performers will experience a smaller drop in performance than the high performers will when the ability heterogeneity turns non-salient. Alternatively, the performance increase is more pronounced in the high performers than in the low performers when the low level of ability heterogeneity becomes salient.

H3: Compared to when the ability heterogeneity is low and salient, the drop in RPI's effectiveness due to a non-salient knowledge of the ability heterogeneity, is lower among the low performers than among the high performers.

3. Method

3.1 Overview

I use a between-subjects experimental design with three conditions: *Homogeneous*, *Heterogeneous*, and *NotSalient*. In *Homogeneous* and *Heterogeneous* conditions, participants are

made aware that they compete in a group of similar or different ability, respectively. In *NotSalient* condition, participants are not informed of any ability information about the group.

The experiment consists of two stages. In the first stage, participants are randomly assigned to groups of four members. Participants compete within-group for 8 minutes by performing a real effort task – "the subtraction slider task". The task is adapted from Chan (2018) and has been found to enhance the effect of incentives on effort intensity (Choi, Clark, and Presslee 2018). During the task, participants are shown a subtraction problem with a corresponding slider initially positioned at 0 along a scrollbar from 0 to 100. The task is to first solve the subtraction problem without the use of calculator or pencil and paper, and then to adjust the slider to the number position that matches the answer to the problem. In each group, participants receive RPI as either "above average" or "below average" based on relative individual performance.

In the second stage, participants are reshuffled and assigned to pairs where they perform the same type of task for another 8 minutes. The manipulation of ability heterogeneity takes effect in this stage. In *Homogeneous* condition, tournament is structured such that participants with the same rank (i.e., low level of ability heterogeneity) from the first stage form a new group in the second stage. In *Heterogeneous* condition, participants with different ranks (i.e., high level of ability heterogeneity) from the first stage form a new group in the second stage. In both conditions, ability heterogeneity is made salient by informing participants whether they are competing with individuals of similar or different abilities in the second stage. In *NotSalient* condition, participants of any ranks from the first stage form a new group in the second stage. Participants' performance is measured by the amount of output they produce, and output is measured by the total number of sliders positioned correctly.

3.2 Design Choices

A few design choices warrant a discussion. First, participants receive a fixed pay rather than a pay-for-performance incentive (e.g., a piece-rate compensation) in the first stage. This is because a fixed pay, unlike any pay-for-performance incentives, does not introduce the wealth effect across the stages.

Second, I choose to form groups of four rather than two in the first stage. Prior literature suggests that individual ability needs not be constant such that the performance may be determined as a process of a stochastic process – that is, in any single event, an underdog may be more able than a favorite (Konrad and Kovenock 2010). The design choice was to convince the participants that RPI accurately reflects their relative ability: in a larger group of four, RPI is associated more with any difference in effort and/or ability than that in noise such as luck.

Third, to allow the full potency of RPI to be carried over into the second stage, I dichotomize individuals between high and low performers, instead of giving specific ranks. Providing specific ranks might lead to nonlinear perceptions of ability heterogeneity, which introduces confounding effects in testing my theory. Also, the dichotomization represents a conservative manipulation as it induces a weaker comparison of performance and thus works against finding the results.

Lastly, I adopt a more conservative manipulation of *NotSalient* condition: giving participants no information about ability heterogeneity. Participants might form some perceptions of ability difference based on their experience in the first stage (i.e., I won/lost, so I'm very likely of higher/lower ability in the next round), and presume those perceptions in the second stage. A stronger manipulation for testing my theory would provide a clean slate of any perceived ability

difference in the second stage, such that participants, if pay any attention to ability heterogeneity, fully consider the 50% chance of competing with a peer of similar ability. In addition, the current manipulation more closely resembles the situation in the real world when companies are silent about tournament structure.

3.3 Procedures

Before participating all participants read through and sign an informed consent form, in accordance with IRB policy. Participants begin the experiment by reading the instructions for the first stage via the computer program. Before proceeding to the tournament, participants take attention-check quizzes and complete a practice round for 60 seconds to get familiar with the subtraction slider task.

Participants are assigned to one of the three conditions based on which session they attend. I run one condition per session and randomly determine which condition is used for each session. All participants receive the same sets of problems in the same order to work on. In the first stage, multiple problems need to be solved within a time limit of 8 minutes. After the first stage, participants receive RPI: either "above average" or "below average" based on their rank with the top two participants in a group above average and the bottom two in a group below average.[3]

After learning their rank for the first stage, participants are reshuffled into new groups based on conditions. In *Homogeneous* condition, a participant is matched up with a new group member who had an identical RPI in the initial group. In *Heterogeneous* condition, a participant

[3] In the event of a tie for the second position, the computer randomly assigns the involved participants into different positions. In the second stage, similarly, the computer randomly picks a winner and the prize is equally split. In the sessions I ran, two pairs of participants tied in *NotSalient* condition and no ties occurred in the other conditions.

is matched up with a new group member who had a different RPI than hers/his in the initial group. In *NotSalient* condition, a participant is matched up with a new group member whose RPI in the initial group is unknown. Participants read the instructions for the second stage before competing for another 8 minutes of the same task. They receive similar feedback information as in the first stage followed by a post-experimental questionnaire. They are paid in cash and dismissed.

3.4 Participants

Participants are recruited from the participant pool of an experimental economics laboratory at a large U.S. public university. Participants are randomly assigned to conditions. All participants receive a fixed reward of $5 by completing the tasks. The winner in each new group in the second stage receives a prize of $10. The experiment is conducted entirely using the z-Tree software (Fischbacher 2007). In total, 180 individuals participated in 7 sessions and formed 90 unique tournaments.[4] Each session took approximately 60 minutes. Participants received a $5 fee, in addition to the payoffs they earned during the experiment. Expected pay of each participant is $15 on average across conditions. In *NotSalient* condition, high performers' expected pay likely is greater than $15 and low performers' likely is less than $15. [5] This only affects and works against finding results for H3.[6]

[4] The size of each session ranges from 20 subjects to 36 subjects. I ran 2 sessions for both *NotSalient* and *Heterogeneous* conditions, and 3 sessions for *Homogeneous* condition due to a limited lab availability before the Thanksgiving Day.

[5] By design, there is a 50% chance a low (high) performer will compete with a high (low) performer in Stage 2. Thus, the low (high) performer's chance of winning is likely lower (higher) than 50%. Ex post, though, low performers reported directionally lower winning expectation (*Performance Expectation in Stage2*, as described in page 19. Mid-point vs. 3.56, $p = 0.11$, one-tailed and untabulated), and high performers reported higher winning expectation (5.6 vs. Mid-point, $p < 0.01$, one-tailed and untabulated).

[6] In *Homogeneous* condition, expected pay is $15 for each participant. In *Heterogeneous* and *NotSalient* conditions, due to the way Stage 2 is structured, high performers' winning chance likely is higher and low performers' is likely lower. Hypothesis test for H1, H2a, and H2b are not affected since average expected pay is constant across conditions. Related to H3, high (low) performers would have worked harder (less hard) in NotSalient than Homogeneous condition with a higher (lower) expected pay, which is against the direction of my prediction.

4. Results

4.1 Participants

Participants, on average, are 19.2 years old, 60.6 percent are female, and are all undergraduate students. There are no significant differences across conditions for gender, age, student status, GPA, country of origin, and undergraduate major (all $p > 0.13$). Participants earned an average of $15.

4.2 Descriptive Results

Table 1 summarizes the descriptive statistics for each condition. Notably, in *Heterogenous* condition, the average level of RPI's effectiveness is much greater than the other conditions (64.4% versus either 3.5% or 17.5%). This is due to the large dispersion (standard deviation = 435.5%) and the exceptionally high maximum (one participant had a 3400% increase in output).

In addition, 10 out of 60 participants reduced performance in *Homogeneous* condition, while the number was 26 out of 60 in *NotSalient* condition and 29 out of 60 in *Heterogeneous* condition.

4.3 Hypotheses Tests

H1 predicts that RPI's effectiveness is higher when the level of ability heterogeneity is low and salient compared to when the level is high and salient. Recall that individual performance is measured by total output during each stage. To capture RPI's effectiveness, I calculate the difference in individual output between the stages, scaled by the output in Stage1 to control for any individual characteristics related to task performance (i.e., RPI's effectiveness = (output in

Stage2-output in Stage1)/output in Stage1). To test H1, I compare RPI's effectiveness between *Homogeneous* and *Heterogeneous* condition. As shown in Panel A of Table 2, RPI's effectiveness was lower when participants have salient knowledge of a low level of ability heterogeneity (a 17.5% increase in performance) than when participants have salient knowledge of a high level of ability heterogeneity (a 64.4% increase in performance), which is inconsistent ($t = -0.83$, $p = 0.41$, two-tailed) with H1.

Based on Tukey's fences (Tukey 1977), four, three, and five participants are considered outliers in *NotSalient*, *Heterogeneous*, and *Homogeneous* conditions, respectively. The distribution of RPI's effectiveness across conditions are statistically equal (all p-value of the Kolmogorov-Smirnov tests are greater than 0.37). I report additional results for all hypotheses tests excluding the outliers.[7] As shown in Panel B of Table 2, excluding the outliers, RPI's effectiveness was higher when the level of ability heterogeneity is high (a 13.8% increase in performance) than when it is low (a 6.8% increase in performance), which is consistent ($t = 1.69$, $p = 0.05$, one-tailed) with H1.

To better understand participants' behavior, I examine their responses to a post-experimental question in which participants indicated, on a seven-point scale with endpoints of "Very unlikely" (1) and "Very likely" (7), the extent to which they expected to win in Stage2 (*Performance Expectation in Stage2*).[8] Participants answered the same question for both *Homogeneous* and *Heterogeneous* conditions in the post-experimental questionnaire. I define participants as high (low) performers if their performance rank in Stage1 is above average (below

[7] If excluding based on three standard deviations of the condition mean, 1, 2, and 0 participants are outliers in *NotSalient*, *Heterogenous*, and *Homogeneous* conditions, respectively. Excluding those observations results in significant support for H1 ($p < 0.01$) and H2a ($p < 0.01$). H2b is a null hypothesis and is supported after excluding the outliers ($p = 0.45$, two-tailed). But, excluding the outliers results in an insignificant support for H3 ($p = 0.16$, two-tailed).

[8] The actual question was "At the beginning of Stage 2, I thought that winning was ___ in Stage 2."

average) among their peers in the same tournament group. As shown in Panel A of Table 2, high performers in both conditions reported greater performance expectation in the second stage than low performers (*Homogeneous*, 4.57 v. 3.30, t = 2.60, p = 0.01, two-tailed; *Heterogeneous*, 5.27 v. 2.83, t = 5.24, p < 0.001, one-tailed). Between the conditions, high performers reported greater performance expectation in *Heterogeneous* than in *Homogeneous* condition (5.27 v. 4.57, t = 1.67, p = 0.05, one-tailed and untabulated), while low performers reported directionally greater performance expectation in *Homogeneous* than in *Heterogeneous* condition (3.30 v. 2.83, t = 0.88, p = 0.19, one-tailed and untabulated). The results suggest that while high performers have greater performance expectation in the second stage compared to low performers regardless of ability heterogeneity knowledge, the difference in the expectation between high and low performers is moderated by such knowledge. This partially supports the underlying theory that the level of ability heterogeneity influenced participants' performance expectations upon receiving RPI. Excluding the outliers does not change the significance. Overall, the results support H1.

H2 predicts that RPI's effectiveness is higher when the ability heterogeneity is not salient than when it is high and salient, but not different than when it is low and salient. To test H2a, I compare RPI's effectiveness across all conditions. As shown in Panel A of Table 3, RPI's effectiveness is greater in *Homogeneous* (a 17.5% increase in performance) than in *NotSalient* condition (a 3.5% increase in performance), which is consistent (F = 7.05, p < 0.01) with H2a. With the outliers, RPI's effectiveness is directionally greater in *Heterogeneous* than in *NotSalient* condition (64.4% vs. 3.5%). But the difference is not statistically significant (F = 1.15, p = 0.29), which is consistent with H2b.

After excluding the outliers, the comparison between *Homogeneous* and *NotSalient* condition remains significant (F = 12.60, p < 0.01) and is consistent with H2a. The difference

between *Heterogeneous* and *NotSalient* condition remains insignificant (F = 2.40, p = 0.12) with a smaller magnitude (from 60.9% to 6.6%).

To test the underlying logic of H2a and H2b, I examine participants' responses to a post-experimental question in which they indicated the extent to which they believed that the peer was of similar task-related ability in Stage2 (*Perceived Heterogeneity*) on a seven-point scale with endpoints of "Disagree" (1) and "Agree" (7)[9]. Inconsistent with the argument behind H2, the results indicate that participants in *NotSalient* condition believed that they were competing in somewhat homogeneous tournaments (mean = 5.10, compared with the midpoint, t = 4.63, p < 0.01, two-tailed and untabulated). This inconsistency might be attributable to the high self-efficacy of the participants in *NotSalient* condition (mean = 4.93, compared with the midpoint, t = 5.16, p < 0.01, two-tailed and untabulated)[10]. Prior studies suggest high self-efficacy leads to low acceptance of feedback (Nease, Mudgett, and Quinones 1999). It is possible that the participants tried to justify their performance by reporting a high perceived homogeneity in order to maintain the self-image.

To further explore whether *Perceived Heterogeneity* explains the relationship between the conditions (*Homogeneous* vs. *NotSalient*) and RPI's effectiveness, I examine a mediation model using the full sample based on Baron and Kenny's (1986) process. The results support a mediation. Besides the direct effect (F = 7.05, p<0.01 two-tailed), the conditions are significantly associated with *Perceived Heterogeneity* (t = 8.01, p<0.01 two-tailed and untabulated), and *Perceived Heterogeneity* explains RPI's effectiveness (t = 1.94, p = 0.05 two-tailed and untabulated) while reducing the impact of the conditions on RPI's effectiveness (coefficient decreases from 0.14 to 0.07). The results suggest the perception of more heterogeneous tournaments in *NotSalient*

[9] The actual question was "In Stage 2, I was competing in a group of similar ability in solving the task."
[10] The measurement of participants' self-efficacy is discussed in section 4.4.4.

condition can be the underlying reason that RPI is less effective compared to *Homogeneous* condition.

Overall, the results support H2a: RPI's effectiveness is lower when the ability heterogeneity is not salient than when it is low and salient. The prediction regarding *Heterogeneous* condition being not different than *NotSalient* condition is also supported.[11]

As illustrated in Panel A of Figure 1, H3 predicts that the drop in RPI's effectiveness between *Homogeneous* and *NotSalient* condition, is lower among the low performers than among the high performers. To test H3, I analyze the interaction term of a two-way ANOVA between RPI (1=above average, 0=below average) and Condition (1=*Homogeneous*, 0=*NotSalient*) on RPI's effectiveness. Inconsistent (Condition*RPI, F = 1.78, p = 0.19) with my prediction, as shown in Panel A of Table 4, there is no significant crossover interaction of Condition and Stage1 Rank. After excluding the outliers, however, the interaction term becomes significant (F = 4.27, p = 0.04). This suggests that the drop in RPI's effectiveness from *Homogeneous* to *NotSalient* condition depends on whether the participants are high performers. The result is significant in supporting H3.

To better understand participants' behavior, I compare their *Performance Expectation in Stage2* with their performance expectation in Stage1[12]. Low performers in *NotSalient* condition reported similar performance expectation in both stages (3.70 v. 3.57, t = -0.28, p = 0.78, two-tailed and untabulated). This suggests that low performers maintain their performance expectation

[11] In *NotSalilent* condition, low (high) performers' *Performance Expectation in Stage2* are insignificantly (significantly) different from the midpoint (mean = 3.57 and 5.60, t = 1.21 and 5.94, p = 0.23 and <0.01, two-tailed and untabulated). The results are partially consistent with my argument that participants discount prior knowledge of their ability.

[12] I examine their responses to a question answered before they receive the ranks in the first stage, in which participants indicated, on a seven-point scale with endpoints of "Very unlikely" (1) and "Very likely" (7), the extent to which they expected to win in Stage1 (*Performance Expectation in Stage1*). The actual question was "I think that winning was ___ in Stage 1."

in Stage2, which, as discussed above, might discourage effort reduction. Also, high performers increased their winning expectation from Stage1 to Stage2 (4.70 v. 5.60, t = 2.25, p = 0.03, two-tailed and untabulated). Consistent with my argument, the higher winning expectation might result in high performers saving effort in Stage2. In comparison, in *Homogeneous* condition, both low and high performers had similar performance expectation between the stages (3.00 v. 3.30 and 4.30 v. 4.57, t = -0.63 and -0.63, p = 0.53 and 0.53, respectively, two-tailed and untabulated). Overall, the results support the underlying theory for H3.

4.4 Supplemental Analysis

4.4.1 Utilities and Disutilties Associated with Tournament Outcomes

In predicting H1, I argue that participants associate nonpecuniary value of winning and losing in Stage2, which vary depending on the level of ability heterogeneity. A participant's focus on either winning or avoiding losing is largely driven by the type of his/her achievement goal. According to achievement goal theory (Elliot and McGregor 2001), individuals situationally pursue two types of goals in competitions: promotional and preventional. In this paper's setting, it is not clear ex ante whether a participant, be the individual a high or low performer, would focus more on any particular achievement goal.[13] Testing participants' achievement goals helps determine whether their nonpecuniary value of tournament outcomes is related to the utility or disutility of the outcomes. Therefore, I examine their responses to a set of three post-experimental questions: the first question is to measure a participant's achievement goal, and the second and

[13] For example, a higher performer may want to win as many times as possible and a low performer may want to win at least once (a maximizer). Alternatively, a high performer may care more about not losing the winner status and a low performer may focus on not losing twice in a row (a satisfier).

third questions are to measure the level of the utilities and disutilities associated with tournament outcomes.

Specifically, in response to the first post-experimental question, participants indicated the extent to which they focused more on winning or on avoiding losing the tournament (*Achievement Goal*) on a seven-point scale with endpoints of "Do better than the other participant" (1) and "Avoid doing poorly in the new group" (7).[14] Next, to capture participants' utilities associated with tournament outcomes, following Krakel (2008) and Moldovanu et al. (2007), I examine participants' responses to a post-experimental question in which they indicated the extent to which they feel positive emotions and/or status recognition about winning the tournament (*Utility*) on a seven-point scale with endpoints of "Not at all" (1) and "To a great degree" (7).[15] Lastly, to capture participants' disutilities associated with tournament outcomes, I examine participants' responses to a post-experimental question in which they indicated the extent to which they feel negative emotions and/or status recognition about losing the tournament (*Disutility*) on a seven-point scale with endpoints of "Not at all" (1) and "To a great degree" (7).[16] *Utility of Ranks* is a participant's perceived value of tournament outcomes conditional on his/her *Achievement Goal*. It equals the participant's reported *Utility* (*Disutility*) if *Achievement Goal* > 4 (< 4). In cases where *Achievement Goal* = 4, *Utility of Ranks* equals the greater response between *Utility* and *Disutility*. As shown in Panel A of Table 2, high performers' *Utility of Ranks* was similar between *Homogeneous* and *Heterogeneous* condition (the means are 4.77 and 5.07, respectively; t = -0.59, p = 0.55, two-tailed); low performers' *Utility of Ranks* was directionally higher in *Homogeneous*

[14] The actual questions were "In Stage 2, did you want to do better than the other participant, or want to avoid doing poorly in the new group? (Adapted from Elliot and McGregor 2001)"
[15] The actual question was "To what degree do you think that winning in Stage 2 would make you feel proud, happy, or fulfilled?"
[16] The actual question was "To what degree do you think that winning in Stage 2 would make you feel proud, happy, or fulfilled?"

than in *Heterogeneous* condition (the means are 5.67 and 5.30, respectively; t = 0.92, p = 0.18, one-tailed).

To further explore participants' *Achievement Goal*, I test a mediation model based on Hayes (2009) process. I examine whether the impact of the conditions (*Homogeneous* vs. *Heterogeneous*) on RPI's effectiveness is mediated by *Achievement Goal*. Using the full sample, the mediating effect is possible. The association between the conditions and Achievement Goal is significant (t = -1.65, p = 0.10 two-tailed and untabulated), *Achievement Goal* reduces the impact between the conditions and RPI's effectiveness (from 0.107 to 0.102) while being insignificantly associated with RPI's effectiveness (t = -0.67, p = 0.50 two-tailed and untabulated). This suggests that participants in *Homogeneous* condition focus more on winning than avoiding losing compared to those in *NotSalient* condition, which might lead to greater performance after receiving RPI.

To support H2a, similarly, participants' responses to the same post-experimental questions indicate that the difference between the value attached to the tournament outcomes in *Homogeneous* and *NotSalient* conditions is consistent with my theory among low performers (the means are 5.66 and 4.53, respectively; t = 2.39, p = 0.01, one-tailed and untabulated). In high performers, however, the difference is not significant (the means are 4.77 and 5.13, respectively; t = -0.74, p = 0.46 two-tailed and untabulated).

In H2b, I argue that the emotions and status-seeking intention aroused by the ability information may encourage spending effort in *Heterogeneous* condition compared to *NotSalient* condition, offsetting effort reduction associated with more pecuniary driving forces such as winning expectancy. Partially consistent with my argument, low performers reported marginally stronger *Utility of Ranks* (5.30 v. 4.53, t = 1.51, p = 0.07, one-tailed and untabulated) and high performers reported similar *Utility of Ranks* (5.07 v. 5.13, t = -0.14, p = 0.89, two-tailed and

untabulated) in *Heterogeneous* than *NotSalient* condition. This suggests the low performers in *Heterogeneous* condition were likely more motivated by non-pecuniary factors than in *NotSalient* condition.

4.4.2 Attribution

To better understand the result of H3, I examine participants' responses to a post-experimental question in which they indicated the extent to which they attributed Stage1 rank to individual differences in effort level or task ability (*Attribution*) on a seven-point scale with endpoints of "Entirely due to differences in individual effort levels" (1) and "Entirely due to differences in individual abilities" (7). The results indicate that, as shown in Panel B of Table 4, high performers attributed their above-average performance in Stage1 to ability (4.93), which is consistent ($t = 3.56$, $p < 0.01$, one-tailed) with part of my theory behind H3. Low performers equally (4.03) attributed their below-average performance to both effort and ability ($t = -0.11$, $p = 0.91$, two-tailed), which does not follow part of my theory behind H3.

4.4.3 Task Attitude

The advanced slider task is meant to test individual behavior generalizable to real corporate tournament settings where intellectual ability is important for career success. To test whether the task speaks to their intellectual ability, I examine their responses to a post-experimental question in which they indicated the extent to which they agreed both intellectual ability and effort are important in doing the advanced slider task (*Task Attitude*) on a seven-point scale with endpoints of "Not at all" (1) and "To a great degree" (7)[17]. Further, before the 60-second practice session, I

[17] The actual question was "To what extent you agree with this statement: In order to solve accurately and quickly the type of subtraction problems given here, one is required to exert effort and possess intellectual ability."

examine their responses to a pre-experimental question in which they indicated the extent to which they thought intellectual ability is important for succeeding in life (*PreAttitude*) on a seven-point scale with endpoints of "Not at all" (1) and "To a great degree" (7).[18] The results indicate that participants in all conditions value intellectual ability in general (the means are 5.67, 5.5, 5.65, respectively in *Homogeneous*, *NotSalient*, and *Heterogeneous* conditions, compared to the mid-point, all $t > 8.25$, all $p < 0.01$, two-tailed and untabulated). Also, they thought intellectual ability was important for performing well in the task in this experiment (the means are 5.13, 5.12, 5.00, respectively in *Homogeneous*, *NotSalient*, and *Heterogeneous* conditions, compared to the mid-point, all $t > 6.01$, all $p < 0.01$, two-tailed and untabulated). In tandem, the reported *PreAttitude* and *Task Attitude* suggest that participants thought their performance in the task reflected their intellectual ability which they believe is important in real life.

4.4.4 Intrinsic Motivation and Self-Efficacy

Both intrinsic motivation and perceived competence (self-efficacy) can potentially affect task performance and make the results discussed above less representative: individuals who are intrinsically motivated to do the task, or with higher self-efficacy, might persistently work hard regardless of feedback or relative ability information. Specifically, prior literature finds that professional athletes do not reduce effort on lagging performance or on inferior ability in tournaments. This is because the productivity of intrinsically motivated individuals is not dependent on social spillovers, whether positive or negative (Guryan et al. 2009).[19] Also,

[18] The actual question was "To what degree do you think that intellectual ability is important for succeeding in life?"

[19] Different from professional athletes who invest in human capital over the course of their life, regular employees arguably are mostly motivated by extrinsic rewards. For instance, in contrast, prior studies using work-like tasks or subjects with regular jobs find peer effects (e.g., Mas and Moretti 2009; Falk and Ichino 2006; Bandiera et al. 2009). Therefore, individuals with high intrinsic motivation are considered less relevant to the context of this paper which is employee tournaments in a regular corporate environment.

individuals with high self-efficacy are less likely to accept feedback (Nease et al. 1999) and more likely to maintain or even increase effort in the face of setbacks (Bandura 1997).

To examine whether the results are affected by their intrinsic motivation in the task, I measure participants' responses to a pre-experimental question in which they indicated the extent to which they enjoyed doing math problems (*Intrinsic*) on a seven-point scale with endpoints of "Not at all" (1) and "To a great degree" (7).[20] The results indicate that participants in *Homogeneous* and *Heterogeneous* conditions were neutral toward the task (the means are 4.00, 4.05, respectively, compared to the mid-point, t = 0.00 and 0.22, p = 1.00 and 0.82, two-tailed and untabulated). In *NotSalient* condition, participants' intrinsic motivation was moderately higher than normal (mean = 4.42, compared to the mid-point, t = 1.86, p = 0.07, two-tailed and untabulated). This indicates participants in *NotSalient* condition might be motivated to maintain effort in the second stage. However, controlling for *Intrinsic* as a covariate does not change the significance of the main results.

Similarly, to examine whether the results are affected by their self-efficacy in the task, I measure participants' responses to a pre-experimental question in which they indicated the extent to which they were able to solve math problems compared to other participants (*Self-Efficacy*) on a seven-point scale with endpoints of "Worse than others" (1) and "Better than others" (7).[21] The results indicate that participants in *Homogeneous* condition perceived neutral relative competence in the task (mean = 4.25, compared to the mid-point, t = 1.27, p = 0.21, two-tailed and untabulated). In *NotSalient* and *Heterogeneous* conditions, participants perceived high relative competence

[20] The actual question was "How much do you think you will enjoy doing the math problems?" The responses were collected before the 60-second practice session.
[21] The actual question was "How well do you think you will do on the math task compared to the other participants in the session?" The responses were collected before the 60-second practice session.

(mean = 4.93 and 4.52, respectively, compared to the mid-point, $t = 5.16$ and 2.82, all $p < 0.01$, two-tailed and untabulated). This indicates self-efficacy might drive the results as a covariate. However, controlling for *Self-Efficacy* as a covariate does not change the significance of the main results of H1, H2, or H3.

Overall, although some participants reported a greater intrinsic motivation or higher self-efficacy, the test results are not significantly affected.

5. Conclusions

I investigate the effects of providing tournament participants with ability heterogeneity information on the effectiveness of relative performance information (RPI). In my experimental setting, the first stage is used to sort out participants' task abilities while the second stage is the tournament that determines the winners. I focus on the effectiveness of RPI provided between the stages. I find that when the ability heterogeneity is low and salient, the performance expectation between high and low performers are closer, resulting in greater performance upon learning the tournament structure. I also find that compared with when the ability heterogeneity is not salient, when the ability heterogeneity is high and salient, RPI's effectiveness is not less negative as winning becomes more necessary or appealing for the low performers. In addition, I find that compared with when the ability heterogeneity is low and salient, when the ability heterogeneity is not salient, high performers are more prone to reducing effort.

My study has several implications. First, the results of my study demonstrate the inter-related nature of tournament structure and information system design choices in designing effective tournaments. The findings extend prior research (Newman and Tafkov 2014) that shows that the negative effect of providing RPI in tournaments, as documented by Hannan et al. (2008),

is not universal to all tournament structures. Second, by showing that the effectiveness of RPI varies based on the level and salience of tournament participants' ability heterogeneity, I highlight that firms need to be conscious about any difference among the employees when RPI is present. Specifically, my results suggest that firms will benefit from emphasizing information about the level of ability heterogeneity to the employees under a homogeneous tournament. The benefit is greater when emphasizing among the more productive employees. Also, if employees can learn RPI via informal sources (Hannan et al. 2013), it is not necessarily the case that firms need to discontinue the tournaments to avoid negative performance effects if the informal information suggests more homogeneous tournaments.

References

Abeler, J., A., Falk, L. Goette, and D. Huffman. 2011. Reference points and effort provision. *American Economic Review* 101: 470–492.

Alicke, M. D., D. S. Vredenburg, M. Hiatt, and O. Govorun. 2001. The "better than myself" effect. *Motivation and Emotion* 25: 7-22.

Altmann, S., A. Falk, and M. Wibral. 2012. Promotions and incentives: The case of multistage elimination tournaments. *Journal of Labor Economics* 30: 149–174.

Akerlof, G., and J. Yellen. 1990. The fair-wage effort hypothesis and unemployment. *Quarterly Journal of Economics* 195: 255–284.

Ashraf, N., O. Bandiera, and S. S. Lee. 2014. Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization* 100: 44–63.

Azmat, G., and N. Iriberri. 2016. The provision of relative performance feedback: an analysis of performance and satisfaction. Journal of Economics & Management Strategy 25 (1): 77–110.

Bandiera, O., I. Barankay, and I. Rasul. 2013. Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association* 11(5): 1079–1114.

Bandura, A. 1997. Self-efficacy: The exercise of control. New York: Freeman.

Barankay, I. 2011. *Rankings and social tournaments: Evidence from a crowd-sourcing experiment.* Working Paper*,* University of Pennsylvania.

Barankay, I. 2012. *Rank incentives: Evidence from a randomized workplace experiment.* Working Paper*,* University of Pennsylvania.

Baron, R. M., and D. A. Kenny. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 51(6): 1173.

Bellemare, C., P. Lepage, and B. Shearer. (2010). Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace. *Labour Economics* 17(1): 276–283.

Berger, L., K. J. Klassen, T. Libby, and A. Webb. 2013. Complacency and giving up across repeated tournaments: Evidence from the field. *Journal of Management Accounting Research* 25(1): 143–167.

Berger, L., T. Libby, and A. Webb. 2018. The effects of tournament horizon and the percentage of winners on social comparisons and performance in multi-period competitions. Accounting, Organizations and Society 64: 1–16.

Berger, J., and D. Pope. 2011. Can losing lead to winning? *Management Science* 57(5): 817-27.

Bhattacharya, H., and S. Dugar. 2012. Status incentives and performance. *Managerial and Decision Economics* 33(7-8): 549–563.

Bidwell, M. J. 2011. Paying more to get less: the effects of external hiring versus internal mobility. *Administrative Science Quarterly* 56 (3): 369-407.

Bonner, S. E., and G. B. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations, and Society* 27: 303–345.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2012. Salience theory of choice under risk. *Quarterly Journal of Economics* 127: 1243-85.

Boudreau, K. J., K. Lakhani, and M. E. Menietti. (2016). Performance responses to competition across skill-levels in rank order tournaments: Field evidence and implications for tournament design. *RAND Journal of Economics* 47(1): 140-165.

Bradler, C., R. Dur, S. Neckermann, and A. Non. (2016). Employee recognition and performance: A field experiment. *Management Science* 62(11): 3085-3099.

Brown, D., L. Ferris, D. Heller, and L. Keeping. 2007. Antecedents and consequences of the frequency of upward and downward social comparison at work. *Organizational Behavior and Human Decision Processes* 102: 59–75.

Brown, J., and F. Gallagher. 1992. Coming to terms with failure: Private self-enhancement and public self-enhancement. *Journal of Experimental Social Psychology* 28: 3–22.

Cai, W., and S. Gallani. 2018. *Does mercy bear richer fruits than strict justice? implicit effects of subjective rewards and penalties in tournaments.* Working paper, Harvard University.

Cardinaels, E., Chen, C. X., and H. Yin. 2018. Leveling the playing field: The selection and motivation effects of tournament prize spread information. *The Accounting Review* 93(4): 127-149.

Casas-Arce, P., and F. Martinez-Jerez. 2009. Relative performance compensation, contests, and dynamic incentives. *Management Science* 55: 1306–1320.

Chan, E. 2018. Promotion, relative performance information, and the Peter Principle. *The Accounting Review* 93(3): 83–103.

Chan, W. 1996. External recruitment versus internal promotion. *Journal of Labor Economics* 14(4): 555-570.

Chen, C., M. Williamson, and F. Zhou. 2012. Reward system design and group creativity: An experimental investigation. *The Accounting Review* 87: 1885–1911.

Choi, J., J. Clark, and A. Presslee. 2018. *Testing the Effect of Incentives on Effort Intensity Using Real-Effort Tasks*. Working paper, University of Wisconsin and University of Illinois.

Delfgaauw, J., R. Dur, A. Non, and W. Verbeke. 2014. Dynamic incentive effects of relative performance pay: a field experiment. *Labour Economics* 28(1): 1-13.

Dohmen, A., A. Falk, K. Fliessbach, U. Sunde, and B. Weber. 2011. Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics* 95: 279–85.

Ederer, F. 2010. Feedback and motivation in dynamic tournaments. *Journal of Economics and Management Strategy* 19: 733–769.

Ederer, F., and E. Fehr. 2017. *Deception and incentives: how dishonesty undermines effort provision*. Working paper, Yale University and University of Zurich.

Elliot, A. J., and H. A. McGregor. 2001. A 2× 2 achievement goal framework. *Journal of personality and social psychology* 80(3): 501.

Eriksson, T., A. Poulsen, and M. C. Villeval. 2009. Feedback and incentives: experimental evidence. *Labour Economics* 16(6):679–688.

Falk, A. and A. Ichino. 2006. Clean evidence on peer effects. *Journal of Labor Economics* 24(1): 39–57.

Fershtman, C., and U. Gneezy. 2011. The tradeoff between performance and quitting in high power tournaments. *Journal of the European Economic Association* 9(2): 318–336.

Festinger, L. 1954. A theory of social comparison. Human Relations 7: 117–40.

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

Freeman, R., and A. Gelber. 2010. Prize structure and information in tournaments: Experimental evidence. *American Economic Journal: Applied Economics* 2(1): 149-164.

Frey, B. S. 2007. Awards as compensation. *European Management Review* 4(1): 6–14.

Garcia, S., and A. Tor. 2007. Rankings, standards, and competition: Task versus scale comparisons. *Organizational Behavior and Human Decision Processes* 102: 95–108.

Gennaioli, N., and A. Shleifer. 2010. What comes to mind. *Quarterly Journal of Economics* 125(4): 1399-1434.

Georganas, S., M. Tonin, and M. Vlassopoulos. 2015. Peer pressure and productivity: The role of observing and being observed. *Journal of Economic Behavior and Organization* 117: 223-232.

Gill, D., Z. Kissova, J. Lee, and V. Prowse. 2018. First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision. *Management Science* 65(2): 459-954.

Gino, F., and B. Staats. 2011. *Driven by Social Comparisons: How Feedback about Coworkers' Effort Influences Individual Productivity*. Working Paper, Harvard Business School.

Gneezy, U., and A. Rustichini. 2004. Gender and competition at a young age. *American Economic Review* 94(2): 377-381.

Greenwald, B. C. 1986. Adverse selection in the labour market. *Review of Economic Studies* 53: 325-347.

Grote, R. C. 2005. *Forced ranking: Making performance management work*. Harvard Business School Press.

Guryan, J., K. Kroft, and M. J. Notowidigdo. 2009. Peer Effects in the workplace: evidence from random groupings in professional golf tournaments. *American Econ. J.: Appl. Econ.* 1 (10): 34–68.

Hamilton, B. H., J. A. Nickerson, and A. Owan. 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy* 111: 465–97.

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

Hannan, R. L., G. McPhee, A. H. Newman, and I. D. Tafkov. 2013. The effects of relative performance information on performance and effort allocation in a multi-task environment. *The Accounting Review* 88(2): 553–576.

Hazels, B., and C. M. Sasse. 2008. Forced ranking: A review. *SAM Advanced Management Journal (Spring)*.

Hayes, A. F. 2009. Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs* **76** (4): 408–420.

Heckman, J. J., T. Jagelka, and T.D. Kautz. 2019. Some contributions of economics to the study of personality (No. w26459). National Bureau of Economic Research.

Heite, J. 2020. Performance in Contests. In: Incentives for Collaboration and Competition. Innovation und Entrepreneurship. Springer Gabler, Wiesbaden.

Kluger, A. N., and A. DeNisi. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119: 254–284.

Konrad, K. A., and D. Kovenock. 2010. Contests with stochastic abilities. *Economic Inquiry* 48(1): 89-103.

Koszegi, B., and M. Rabin. 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics* 121 (2006): 1133-1166.

Kräkel, M. 2008. Emotions in tournaments. *Journal of Economic Behavior and Organization* 67(1): 204-214.

Kuhnen, C.M., and A. Tymula. 2012. Feedback, self-esteem, and performance in organizations. *Management Science* 58 (1): 94-113.

Kulik, C. T., and M. L. Ambrose. 1992. Personal and situational determinants of referent choice. *Academy of Management review* 17(2): 212-237.

Kwoh, J. 2012. ‘‘Rank and yank’’ retains vocal fans. Wall Street Journal (January 31). https://www.wsj.com/articles/SB10001424052970203363504577186970064375222

Lambert, R. A. 2007. Agency theory and management accounting. In *Handbook of Management Accounting Research*, Vol. 1, edited by C. Chapman, A. Hopwood, and M. Shields. Oxford, U.K.: Elsevier.

Lazarus, R. 1991. *Emotion and Adaptation*. New York, NY: Oxford University Press.

Lazear, E., and S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89 (5): 841–864.

Lise, J., C. Meghir, and J. Robin. 2016. Matching, sorting and wages. *Review of Economic Dynamics* 19:63-87.

Martin, R. 2000. ‘‘Can I do X?’’ Using the proxy comparison model to predict performance. In *Handbook of Social Comparison: Theory and Research*, edited by J. Suls and L. Wheeler, 67–80. New York, NY: Kluwer Academic/Plenum Publishers.

Mas, A., and E. Moretti. 2009. Peers at work. *American Economic Review* 99(1):112–45.

Milgrom, P., and J. Roberts. 1992. Economics, organization and management. Englewood Cliffs, NJ: Prentice Hall.

Moldovanu, B., A. Sela, and X. Shi. 2007. Contests for status. *Journal of Political Economy* 115:338–63.

Moore, D. A., and P. J. Healy. 2008. The trouble with overconfidence. *Psychological Review* 115(2): 502–517.

Muller, W., and A. Schotter. Workaholics ad dropouts in organizations. *Journal of the European Economic Association* 8 (2009): 717-743.

Nease, A. A., B. O. Mudgett, and M. A. Quiñones. 1999. Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. *Journal of applied psychology* 84(5): 806.

Newman, A. H., and I. D. Tafkov. 2014. Relative performance information in tournaments with different prize structures. *Accounting, Organizations and Society* 39(5): 348–361.

Nordstrom, R., P. Lorenzi, and R. Hall. 1990. A Review of public posting of performance feedback. *Journal of Organizational Behavior Management* 11: 101–123.

Parco, J. E., A. Rapoport, and W. Amaldoss. 2005. Two-stage contests with budget constraints: an experimental study. *Journal of Mathematical Psychology* 49:320-38.

Perez-Truglia, R. 2019. *The effects of income transparency on well-being: evidence from a natural experiment.* Working paper, University of California Los Angeles.

Piazza, A., and F. Castellucci. 2014. Status in organization and management theory. *Journal of Management* 40(1): 287-315.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37 (1): 7–63.

Rosaz, J., R. Slonim, and M. C. Villeval. 2012. Quitting and peer effects at work. *IZA Discussion Paper 6475.*

Sheremeta, R. M. 2010. Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior* 68(2): 731–747.

Smith, R. 2000. Assimilative and contrastive emotional reactions to upward and downward social comparisons. In *Handbook of Social Comparison: Theory and Research*, edited by J. Suls and L. Wheeler, 173–200. New York, NY: Kluwer Academic/Plenum Publishers.

Smither, J. W., M. London, and R. R. Reilly. 2005. Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel psychology* 58(1): 33-66.

Tafkov, I. D. 2013. Private and public relative performance information under different compensation contracts. *The Accounting Review* 88 (1): 327–350.

Taylor, S. E., and S. C. Thompson. 1982. Stalking the elusive 'vividness' effect. *Psychological Review* 89 (March): 155-181.

Tukey, J. W. 1977. Exploratory data analysis (Vol. 2). In: Reading.

Waldman, M. 1984. Job assignments, signaling, and efficiency. *RAND Journal of Economics* 15: 255-267.

TABLE 1

Descriptive Statistics of Performance and RPI's Effectiveness

	Stage-1 Output	Stage-2 Output	RPI's Effectiveness[22]	25% RPI's Effectiveness	75% RPI's Effectiveness	Minimum RPI's Effectiveness	Maximum RPI's Effectiveness
NotSalient **(n=60)**	27.9 (14.2)	28.1 (14.1)	3.5% (29.7%)	-13.7%	15.4%	-66.7%	110.0%
Heterogeneous **(n=60)**	28.3 (12.4)	29.7 (10.9)	64.4% (435.5%)	-9.3%	24.5%	-81.3%	3400.0%
Homogeneous **(n=60)**	29.8 (13.4)	34.2 (14.0)	17.5% (27.5%)	0.0%	30.9%	-57.1%	95.7%

[22] RPI's effectiveness is the difference of individual output between Stage2 and Stage1, scaled by Stage1 output.

TABLE 2

RPI's Effectiveness When Ability Heterogeneity Is Salient (H1)

Panel A: Full Sample

	Homogeneous [n = 60]	**Heterogeneous [n = 60]**	**(t-test)**
RPI's effectiveness	17.5% (0.27)	64.4% (4.35)	Diff. = -46.9% (t = -0.83, p = 0.41)

	Homogeneous		
	High Performers [n = 30]	**Low Performers [n = 30]**	**(t-test)**
Performance Expectation in Stage2	4.57 (1.67)	3.30 (2.12)	Diff. = 1.27 (t = 2.60, p = 0.01)

	Heterogeneous		
	High Performers [n = 30]	**Low Performers [n = 30]**	**(t-test)**
Performance Expectation in Stage2	5.27 (1.53)	2.83 (1.92)	Diff. = 2.00 (t = 5.24, p < 0.01)

	High Performers		
	Homogeneous [n = 30]	**Heterogeneous [n = 30]**	**(t-test)**
Utility of Ranks[23]	4.77 (0.74)	5.07 (0.99)	Diff. = -0.30 (t = -0.59, p = 0.55)

	Low Performers		
	Homogeneous [n = 30]	**Heterogeneous [n = 30]**	**(t-test)**
Utility of Ranks	5.67 (1.01)	5.30 (1.17)	Diff. = 0.37 (t = 0.92, p = 0.18)

[23] *Utility of Ranks* is the reported *Utility* or *Disutility* conditional on participants' *Achievement Goal*.

Panel B: Excluding Outliers

	Homogeneous [n = 55]	**Heterogeneous [n = 57]**	**(t-test)**
RPI's effectiveness	13.8% (0.19)	6.8% (0.24)	Diff. = 7.0% (t = 1.69, p = 0.05)

	Homogeneous		
	High Performers [n = 30]	**Low Performers [n = 25]**	**(t-test)**
Performance Expectation in Stage2	4.57 (1.67)	3.04 (1.99)	Diff. = 1.53 (t = 3.01, p < 0.01)

	Heterogeneous		
	High Performers [n = 30]	**Low Performers [n = 27]**	**(t-test)**
Performance Expectation in Stage2	5.27 (1.53)	2.75 (1.79)	Diff. = 2.52 (t = 5.40, p < 0.01)

	High Performers		
	Homogeneous [n = 30]	**Heterogeneous [n = 30]**	**(t-test)**
Utility of Ranks	4.77 (0.74)	5.07 (0.99)	Diff. = -0.30 (t = -0.59, p = 0.55)

	Low Performers		
	Homogeneous [n = 25]	**Heterogeneous [n = 27]**	**(t-test)**
Utility of Ranks	5.72 (1.34)	5.19 (1.68)	Diff. = 0.53 (t = 1.25, p = 0.11)

TABLE 3

RPI's Effectiveness Depending on the Salience of Ability Heterogeneity (H2a and H2b)

Panel A: Full Sample

	Homogeneous [n = 60]	**NotSalient [n = 60]**	**(ANOVA)**
RPI's effectiveness	17.5% (0.27)	3.5% (0.30)	Diff. = 13.9% (F = 7.05, p < 0.01)

	Heterogeneous [n = 60]	**NotSalient [n = 60]**	**(ANOVA)**
RPI's effectiveness	64.4% (4.35)	3.5% (0.30)	Diff. = 60.9% (F = 1.15, p = 0.29)

Panel B: Excluding Outliers

	Homogeneous [n = 55]	**NotSalient [n = 56]**	**(ANOVA)**
RPI's effectiveness	13.8% (0.19)	0.2% (0.21)	Diff. = 13.6% (F = 12.6, p < 0.01)

	Heterogeneous [n = 57]	**NotSalient [n = 56]**	**(ANOVA)**
RPI's effectiveness	6.8% (0.24)	0.2% (0.21)	Diff. = 6.6% (F = 2.40, p = 0.12)

TABLE 4
Drop in RPI's Effectiveness (H3)

Panel A: Two-Way ANOVA, Full Sample

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Condition	1	0.5869	0.5869	7.57	0.0069
Stage1 Rank	1	0.6916	0.6916	8.92	0.0034
Condition*Stage1 Rank	1	0.1377	0.1377	1.78	0.1851

Two-Way ANOVA, Excluding Outliers

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Condition	1	0.4719	0.4719	12.41	0.0006
Stage1 Rank	1	0.2159	0.2159	5.68	0.0189
Condition*Stage1 Rank	1	0.1625	0.1625	4.27	0.0412

Panel B: NotSalient, Full Sample

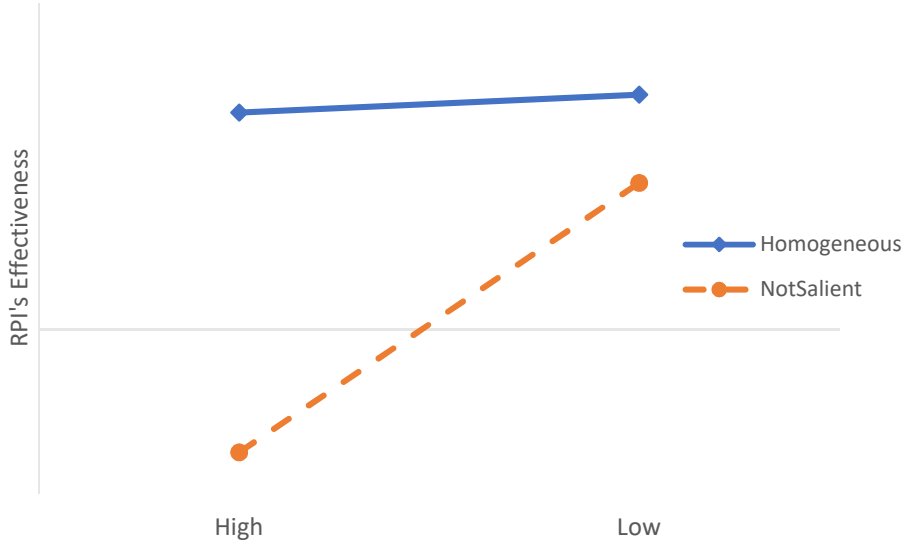
	High Performers [n = 30]	**(t test)**	**Low Performers [n = 30]**	**(t test)**
Attribution	4.93 (1.41)	Diff. = 0.93 (t = 3.56, p < 0.01)	4.03 (1.58)	Diff. = 0.03 (t = -0.11, p = 0.91)

NotSalient, Excluding Outliers

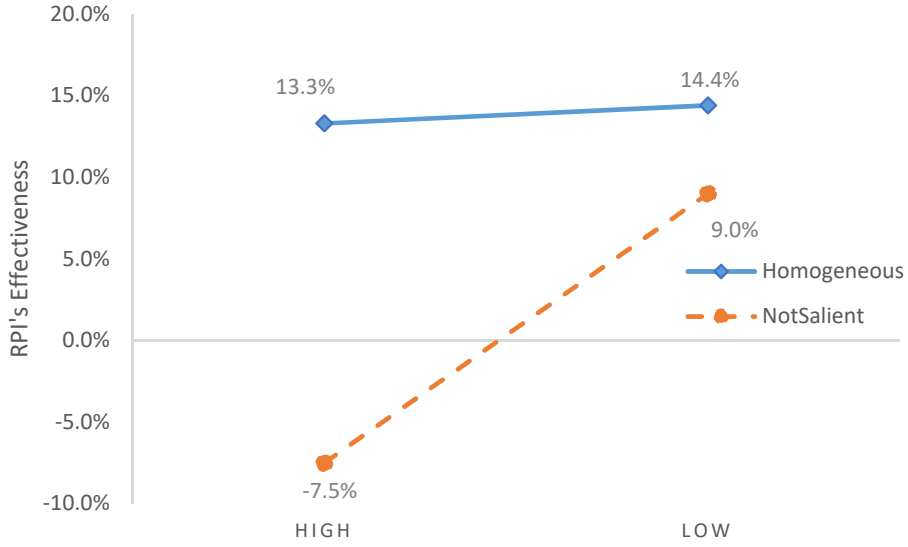
	High Performers [n = 30]	**(t test)**	**Low Performers [n = 26]**	**(t test)**
Attribution	4.93 (1.41)	Diff. = 0.93 (t = 3.56, p < 0.01)	4.15 (1.59)	Diff. = 0.15 (t = 0.49, p = 0.63)

FIGURE 1
Drop in RPI's Effectiveness (H3)

Panel A: Predicted



Panel B: Results



Instructions

General

Welcome!

You are about to participate in a study on problem solving. Please read the instructions carefully. You need to understand these instructions to make money today. In total, this session should last no longer than 75 minutes.

No one will be able to associate your responses in today's study with you personally. Responses will only be attributable to your unique participation ID. You chose your ID by selecting where you sat in the room. As such, there is no way to associate your name with a particular participation ID.

You are prohibited from using a cell phone, a pen / pencil, paper, calculator, or any other outside materials during this session.

If you have any questions at any time during today's session, please raise your hand. The administrator will answer your question in private. Please do not communicate with anyone other than the administrator after this point.

Overview of the Study

This study consists of two stages. Your task is to solve a series of problems in each stage.
You will earn income in an experimental currency called *Lira*. At the conclusion of today's session, the *Lira* you earn will be converted to dollars at the rate of $1 for 60 *Lira*. Therefore, the more *Lira* you earn, the more money you will earn. The amount of money you earn will be paid out in cash at the end of the session.

Stage 1

At the beginning of Stage 1, you will be **randomly** matched with three other participants to form an initial group of **four (4)**. You will not learn the identity of the other members of your group and they will not learn your identity.

In Stage 1, your task is to solve a series of problems. The computer screen will display one problem at a time. **After solving the current problem, you must click the "SUBMIT" button, to work on the next problem**.

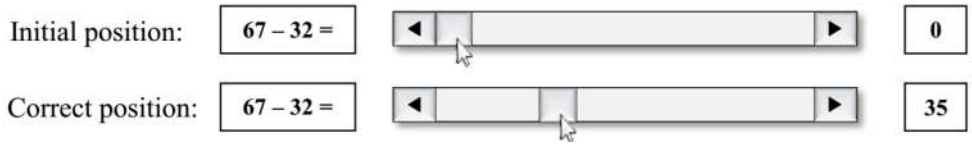
Stage 1 will last 480 seconds (8 minutes). To help you keep track of your time, a clock is displayed on all screens during the stage. The clock counts down from 480 seconds (8 minutes) and starts at the beginning of the stage.

All participants will work on the **exact same set of problems in the exact same order**.

Task Description (Stage 1)

In this stage, you will perform a 'math slider task.' The computer screen will display a few sliders along with some subtraction problems, with each slider initially positioned at 0 along a scrollbar from 0 to 100. **Your task is to first solve a subtraction problem and then adjust the slider to the number position that matches your answer to the problem.**

Below is a screenshot sample of the task.



In this sample task, for instance, a subtraction problem of '67-32=' is displayed to the left of the scrollbar. The number displayed to the right of the scrollbar indicates the current number position of the slider, which changes as you move the slider.

To complete the task, you need to first solve in your head the subtraction problem. Assume your solution to the problem is 67-32=35, you must then drag the slider from the initial position of 0 to the position of 35 using your computer mouse.

Math problems are often used in standardized tests administered for determining college and graduate school entrance. In order to solve accurately and quickly the type of subtraction problems given here, one is required to exert **effort** and possess **intellectual ability.**

[Question A1, A2, and A3 about here.]

Practice Period

You will start with a practice period. The practice period will allow you to become familiar with the subtraction problems and the program. You will NOT be paid for the practice round.

You will have 60 seconds to perform this task. A timer will show in the top right hand corner of the screen.

Please click the 'START TASK' button to begin the task.

Compensation (Stage 1)

You will earn 150 *Lira* for completing the task. **Your performance in this stage may affect some aspects of the next stage**.

Feedback (Stage 1)

At the end of the stage, you will learn how many problems you solved correctly in Stage 1.

In addition, at the end of Stage 1, **you will also learn whether your performance (number of correctly solved subtraction problems) was BETTER or WORSE than at least two members in your group. Specifically, a message will appear on the computer screen which ranks (Above Average or Below Average) your performance among the four participants in your group**.

Quiz 1

Before we move on, you need to answer ALL the following questions correctly. If you encounter difficulties in answering questions, please raise your hand and a facilitator will assist you.

1. To correctly solve a problem, you must both find the right answer and move the slider to the right position.
   True
   False

2. Your rank reflects your performance compared to at least two other participants on your group.
   True
   False

3. You are matched with three other participants to form an initial group of four in Stage 1.
   True
   False
4. Participants will work on different sets of problems in Stage 1.
   True
   False

5. You earn 150 *Lira* for completing the task.
   True
   False

(Transition)

You have correctly answered all the questions in Quiz 1. The actual task (Stage 1) is about to start.

Stage 2

(Homogeneous Condition) In this new stage, you will be assigned to **a NEW group of TWO** participants. **Importantly, the other participant in your new group had the *same* rank as yours**

in Stage 1. That is, for example, if your performance rank was **Above Average** in Stage 1, the other participant assigned to your group also had performance rank **Above Average** in Stage 1. If your performance rank was **Below Average** in Stage 1, the other participant assigned to your group also had performance rank **Below Average** in Stage 1.

You will not learn the identity of the other member of your group and he/she will not learn your identity.

(*Heterogeneous Condition*) In this new stage, you will be assigned to **a NEW group of TWO** participants. **Importantly, the participant in your new group had a *different* rank than yours in Stage 1**. That is, for example, if your performance rank was **Above Average** in Stage 1, the other participant assigned to your group had performance rank **Below Average** in Stage 1. If your performance rank was **Below Average** in Stage 1, the other participant assigned to your group had performance rank **Above Average** in Stage 1.

You will not learn the identity of the other member of your group and he/she will not learn your identity.

(Non-Salient Condition) In this new stage, you will be assigned to **a NEW group of TWO** participants.

You will not learn the identity of the other member of your group and he/she will not learn your identity.

Task Description (Stage 2)

After being assigned to your new group, once again you will perform the math slider task for 480 seconds (8 minutes). To help you keep track of your time, a clock is displayed on all screens during the stage. The clock counts down from 480 seconds (8 minutes) and starts at the beginning of the stage.

All participants will work on the **exact same NEW set of problems in the exact same order. The difficulty level of the new problems remains the same** as in the previous stage.

Feedback (Stage 2)

By the end of the stage, you will also learn how many problems you solved correctly in Stage 2. **In addition, a message will appear on the computer screen which shows your relative performance in Stage 2 among the two participants in your new group (Won or Lost)**.

Compensation (Stage 2)

Your output in the previous stage does not count towards your output in this stage. Recall that you have already earned 150 *Lira* for completing the previous stage. In this stage, each participant receives a salary of 150 *Lira* for completing the task regardless of accuracy. **Moreover,**

the Winner of each new group in Stage 2 will receive a bonus of 600 *Lira*. In the event of a tie, the bonus amount will be split between both participants.

At the end of the session, the *Lira* from both stages will be added together to determine the total *Lira* that you earned. The amount will be converted to dollars and the resulting amount will be paid to you in cash at the end of today's session.

Quiz 2

Before we move on, you need to answer ALL the following questions correctly. If you encounter difficulties in answering questions, please raise your hand and a facilitator will assist you.

1. *(Homogeneous condition)* The other participant in your new group had the same performance rank as yours in Stage 1.
   True
   False

   (Heterogeneous condition) The other participant in your new group had a different performance rank as yours in Stage 1.
   True
   False

2. If a participant's rank were Above Average in Stage 1 and Last in Stage 2, he/she will receive a bonus of 600 *Lira* on top of the salary.
   True
   False
3. If a participant's rank were Below Average in Stage 1 and First in Stage 2, he/she will receive a bonus of 600 *Lira* on top of the salary.
   True
   False

4. You are paired with another participant to form a new group of two in Stage 2.
   True
   False

5. Participants will work on the exact same **NEW** set of problems in the exact same order in Stage 2.
   True
   False

(Transition)

You have correctly answered all the questions in Quiz 2. The actual task (Stage 2) is about to start.

The following instruments will be displayed on the subjects' screens during the actual tasks.

Task – Stage 1

You will now begin Stage 1. You will have 480 seconds (8 minutes) to perform the math slider task during Stage 1.

At the end of the 480 seconds (8 minutes), you will learn the number of problems correctly solved **and your rank**.

You are assigned to a group of **four (4)** participants.

Please click the 'START TASK' button to begin the task.

[Question B about here]

Feedback – Stage 1

The number of problems you correctly solved in Stage 1:_____.
Your rank in Stage 1: **(Above Average / Below Average).**
Task – Stage 2

You will now begin Stage 2. You will have 480 seconds (8 minutes) to perform the math slider task during Stage 2. At the end of the 480 seconds (8 minutes), you will learn the number of problems correctly solved **and your rank**.

You are assigned to a **new** group of **two (2)** participants.

(Homogeneous/Heterogeneous condition)
Your performance in Stage 1 was **(Above Average / Below Average).**
The other member on the new group was **(Above Average / Below Average)** in his/her initial group in Stage 1.
Please click the 'START TASK' button to begin the task.

(Non-Salient condition)
Your performance in Stage 1 was **(Above Average / Below Average).**
Please click the 'START TASK' button to begin the task.

Feedback – Stage 2

You correctly solved _____ problems in Stage 2.
You **(Won/Lost)** in Stage 2.

Pre-Experimental Questionnaires

A1. (To be collected before the practice round) Construct: intrinsic motivation

How much do you think you will enjoy doing the math problems?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not at all			Somewhat			Very much

A2. (To be collected before the practice round) Construct: Self-efficacy

How well do you think you will do on the math task compared to the other participants in the session?

1	**2**	**3**	**4**	**5**	**6**	**7**
Worse than others			About the same as others			Better than others

A3. Construct: PreAttitude

To what degree do you think that intellectual ability is important for succeeding in life?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not at all			To a moderate degree			To a great degree

B. (To be collected before participants receive their rank in Stage 1) Construct: performance expectation in Stage 1

I think that winning was ___ in Stage 1.

1	**2**	**3**	**4**	**5**	**6**	**7**
Very Unlikely						Very Likely

Post-Experimental Questionnaires

NS1. Non-Salient condition. Construct: perceived ability heterogeneity

In Stage 2, I was competing in a group of similar ability in solving the task.

1	**2**	**3**	**4**	**5**	**6**	**7**
Disagree						Agree

1./NS2. Construct: performance expectation in Stage 2

At the beginning of Stage 2, I thought that winning was _____.

1	2	3	4	5	6	7
Very Unlikely						Very Likely

NS3. Non-Salient Condition. Construct: attribution

Do you think that your rank in Stage 1 (being either Above Average or Below Average) was due more to differences in individual effort levels or more to differences in individual abilities?

1	2	3	4	5	6	7
Entirely due to differences in individual effort levels			Equally due to differences in effort and ability			Entirely due to differences in individual abilities

2./NS4. Construct: achievement goal

In Stage 2, did you want to do better than the other participant, or want to avoid doing poorly in the new group? (Adapted from Elliot and McGregor 2001)

1	2	3	4	5	6	7
Do better than the other participant						Avoid doing poorly in the new group

3./NS5. Construct: Utility

To what degree do you think that winning in Stage 2 would make you feel proud, happy, or fulfilled?

1	2	3	4	5	6	7
Not at all			To a moderate degree			To a great degree

4./NS6. Construct: Disutility

To what degree do you think that losing in Stage 2 would make you feel ashamed, upset, or unfulfilled?

1	2	3	4	5	6	7
Not at all			To a moderate degree			To a great degree

5./NS7. Construct: attitude toward the math slider task

To what extent you agree with this statement: In order to solve accurately and quickly the type of subtraction problems given here, one is required to exert effort and possess intellectual ability.

1	**2**	**3**	**4**	**5**	**6**	**7**
Not at all			To a moderate degree			To a great degree

Demographic Questions

1. What is your gender?

2. What is your age?

3. What is your education level?

4. What is your major?

5. What is your GPA?

6. What country are you originally from?