Invoking Social Comparison to Improve Performance by Ranking Employees: The Moderating Effects of Public Ranking, Rank Pay, and Individual Risk Attitude

C. Bram Cadsby [1]
University of Guelph
bcadsby@uoguelph.ca

Fei Song [2]
Ryerson University
fsong@ryerson.ca

Jim Engle-Warnick
McGill University
Jim.engle-warnick@mcgill.com

Tony Fang
Memorial University
ftang06@gmail.com

[1] Corresponding Author. Department of Economics and Finance, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada.
[2] The two first authors, listed alphabetically, contributed equally.

Abstract

Inspired by social-comparison theory, we examine the effectiveness of relative performance ranking as an inherent incentive mechanism to enhance productivity, specifically testing the possibility that the effect is moderated by two features of the feedback design: private/public ranking (whether ranking information was released privately to each individual or announced publicly to all) and fixed/rank pay (whether pay is fixed or positively and monotonically based on rank). Furthermore, generalizing from the theoretical and experimental literature on bids in contests, we introduce individual attitudes toward risk as a potential moderator associating stronger incentive effects of feedback with those who are more tolerant toward risk. We test our hypotheses through a real-effort lab experiment with university students and a companion lab-in-field experiment with full-time employees. We empirically demonstrate both the positive effect of performance-ranking feedback on performance for those sufficiently tolerant of risk and also the statistically significant and practically important moderating effect of risk attitude for both students and factory employees. However, we also find important differences between the two populations regarding public ranking, rank pay and risk attitudes, illustrating the limitations of examining workplace phenomena using exclusively student populations.

Keywords: relative performance ranking; social comparison; public ranking; rank pay; risk aversion; lab-in-field experiment.

JEL Codes: M52, C91, C93.

PsycINFO Classification codes: 3620, 3660.

Highlights

- We study relative performance ranking as an inherent incentive for productivity.
- Public ranking, rank pay and individual risk attitude are examined as moderators.
- We conduct a real effort experiment on university students and full-time
- employees.
- Data corroborate a positive effect of performance ranking on performance.
- The size of this effect is greater for those who are less risk-averse.

Many organizations utilize information about an employee's performance rank relative to his or her peers as key performance feedback. For example, car dealerships and realty companies routinely communicate individual sales rankings to their employees. Some companies also display such rankings on a public Wall of Honor. Students are often keen to find out how they rank within their cohort, and professors are frequently asked to provide such rankings in reference letters supporting student applications for further study. Rankings are omnipresent in the market place as well. There is a plethora of websites devoted to ranking restaurants, hotels and resorts, as well as teachers, professors, dentists and medical doctors. Universities are regularly ranked both in general and by discipline, and such rankings play an important role when students consider where to study.

Can the knowledge that one will receive ranking information on one's performance relative to others motivate better performance, and does such motivation depend on whether such rankings are linked to financial rewards, and/or conveyed privately or publicly? Social-comparison theory (Festinger, 1954), which focuses on the internal drive of individuals to rank well relative to others, and status theory (Frank, 1985), which focuses on people's need for social recognition or reputation among peers both suggest circumstances under which privately or publicly conveyed ranking information, even when decoupled from financial rewards, could be an important motivator.

Our study begins with a simple baseline environment in which employees are paid a fixed salary unrelated to performance, and are provided with no information regarding their performance rank. It then examines how the provision of such ranking information affects performance, focusing on three important potential moderating contingencies: whether the information is provided privately to each individual employee or announced publicly to all (public ranking); whether ranking is merely information on one's relative performance or linked to one's pay (rank pay); and the attitude toward risk of the individual employee receiving the ranking information (risk attitude).

The first of these moderators, public ranking, has been studied before in a fixed-salary context. Both Hannan et al. (2012) and Tafkov (2013) find that when information is provided publicly in a fixed-salary environment, the positive effects on performance are stronger than when it is provided only privately. However, these studies as well as others that examine either private (Kuhnen & Tymula, 2012) or public (Charness et al., 2014) provision under fixed wages all use students as participants. In other related studies that examine the effect of performance-ranking information on performance under piece rates, results appear to be quite different for students than for employees (e.g., see Hannan et al., 2008 and Tafkov, 2013 for positive effects among students and Barankay, 2011, 2012 for negative effects among employees[1]). Though there are a variety of possible explanations for these differences, they nonetheless point to the important need to examine the robustness of the Hannan et al. (2012) and Tafkov (2013) results for employees when compared to students. For this reason, we conduct two studies: a lab-in-field study using factory workers and a traditional laboratory experiment using students from the same geographical region.

The second proposed moderator is whether or not rank is linked to pay. When such a link exists, the calculation and provision of ranking information is no longer merely a psychological incentive, but adds a financial incentive to improve one's rank. We consider a rank-based pay structure, rather than a winner-take-all contest. With rank-based pay, pay is positively and monotonically based on performance rank. Thus, in contrast to a winner-take-all tournament where with heterogeneous participants, those ranked low may lose heart and stop exerting effort (see Dechenaux et al., 2015 for a discussion of the literature concerning this "discouragement effect"), every participant has a financial incentive to work diligently to achieve as high a rank as possible. There is empirical literature documenting that multiple prizes given to more participants may indeed be more motivating than a single large prize, which many participants may have little chance of winning (e.g., Freeman & Gelber, 2010). This is also consistent with Dutcher et al.

[1] In contrast, So et al. (2017) find no significant effects of feedback on rank in two-person student groups under piece rates while Blanes i Vidal & Nossol (2011) find positive effects of private ranking information under piece rates among employees.

(2015), who show that incorporating incentives both at the top and at the bottom of the ranking induces higher effort than incentives only to achieve the top position, especially in larger groups. Linking financial rewards throughout the ranking distribution to ranking information that could otherwise have a motivating effect only through social comparison tests the strength of mere social comparison as a motivator. If social comparison of relative performance alone is a strong enough incentive, it may elicit maximum effort, leaving little or no room for additional financial incentives to have any beneficial effect on performance.

The third proposed moderator, risk attitude, is different from the others in that it is based on the individual preferences of the recipients of ranking information rather than the policy choices made by the organization for which those individuals are working. We argue that people can react differently to the same information based on individual personality factors, and we begin in this study with the examination of one such factor that seems especially relevant when increased effort, though positively related to one's expected rank, interacts with a number of random factors to produce one's actual rank. A risk-averse person may care deeply about his or her rank as per social-comparison theory, but at the same time be reticent to suffer the psychological costs of exerting a lot of effort when the results of the exertion are so risky. This may be the case whether the utility of being ranked high is psychologically based on shame and pride alone or also financially based on rank-linked pay.

In addition to conducting an experimental laboratory study with student participants, we conducted a lab-in-field experiment among full-time employees in an actual work place. Our central behavioral prediction, hypothesizing a positive effect of performance ranking information on performance, was corroborated by both student and employee data. However, we also found important differences between the two populations regarding public ranking, rank pay and risk attitude. These observations underlie our contributions to understanding the applicability and limitations of social-comparison theory as a framework for understanding the impact of feedback regarding performance rank on task performance itself.

METHODS

Research Design

We conducted a real-effort lab experiment with university students and a companion artefactual lab-in-field experiment with laboratory controls (Harrison & List, 2004) using full-time factory employees in China. In our two experiments, participants repeatedly perform a number-addition task, a standard real-effort task in the literature (e.g., Cadsby et al., 2013; Niederle & Vesterlund, 2007). Each number was drawn from the uniform distribution using a computerized random-number generator. Participants solved these arithmetic problems over eight experimental rounds, with each round lasting three minutes. Individual performance in a given round was defined as the number of correct solutions for that round. This particular task is ideal as it is similar in nature to many routine jobs in real organizational life. Moreover, while both full-time employees and university students can accomplish it, it may be taxing under the pressure of a time constraint. We conducted the experiment using paper-and-pencil so that computer proficiency was not an issue for any participant.

There were five treatments, all of which consisted of eight rounds of play. In each of the five treatments, participants experienced the control condition of fixed pay and no ranking information in both the first two and the last two rounds. Thus, the treatments differed only during the middle four rounds. In the control treatment (hereafter denoted C), participants were exposed to the fixed-pay with no ranking information condition during the middle four rounds as well. Specifically, each participant earned ¥6.00 per round regardless of his/her performance and received no feedback on that performance. Performance under the control condition reflects the effort levels chosen by participants in the absence of either rank feedback inducing social comparison or performance pay. If exerting any effort at all were burdensome, participants would choose to do nothing in such circumstances. However, we expect this would rarely happen because most people prefer some activity to doing nothing. The effort exerted thus represents each person’s intrinsic preferences regarding the most comfortable level of activity when that activity has no extrinsic consequences.

Since learning by doing may occur, we expect this level of effort to result in better performance over time and in particular to improved performance during the middle four than during the first two rounds even in the control treatment. We use this control treatment as a within-person baseline to compare how much performance improves under the varying public/private ranking and rank/fixed pay conditions. In particular, our focus is to investigate whether the four experimental treatments, which differ from the control treatment only during the middle four rounds, affect the within-person **change** in performance from the first two to the middle four rounds. The resulting difference-in-difference analysis enables us to avoid confounding learning effects with the effects of the various experimental treatments. Note that there is no reason such learning from experience should be related to risk attitude. Therefore, the moderating effect of individual risk attitudes applies only to the four experimental treatments, and not to the control.

We also employ a similar difference-in-difference design to examine the within-person **change** in performance from the middle four to the last two rounds. This enables us to explore whether any improvements due to the experimental conditions imposed during the middle four rounds carry over to the last two rounds or are reversed when those conditions are removed.

For the experimental treatments, we utilized a 2×2 factorial design that manipulated private/public ranking (private ranking means privately revealed information while public ranking means publicly revealed information) and fixed/rank pay (fixed pay was ¥6.00 per round while rank pay was rank-based pay) for the middle four rounds. Accordingly, our four experimental treatments are: Private/Fixed Pay, Public/Fixed Pay, Private/ Rank Pay, and Public/Rank Pay). In the Private/Fixed Pay treatment, each participant earned ¥6.00 per round regardless of his/her performance just as in the control. However, each person's performance was ranked relative to the other participants, and each participant was privately informed of his/her own ranking on a small piece of paper. If there were ties, the tied participants received identical rankings. For example, if three participants were tied for second behind one participant who had the highest ranking, the tied participants would all receive a ranking of second. The participant ranked immediately below them would then be ranked fifth. In the Public/Fixed Pay treatment, each participant earned ¥6.00 per round regardless of his/her performance just as in the control and the Private/Fixed Pay treatments. Once again, each person's performance was ranked relative to the other participants. However, in contrast to the Private/Fixed Pay treatment, the rankings of all participants were announced publicly by reading out the names of the participants and their ranking in the order of their performance rank. Thus, rankings became public information in that everyone knew each other's ranking and everyone knew that all others knew everyone's ranking. In the Private/Rank Pay treatment, each performance was ranked and reported privately, and each participant was paid based on his/her relative performance ranking according to the pre-announced schedule discussed in the next paragraph. In the Public/Rank Pay treatment, performance was ranked and all rankings were publicly announced just as in the Public/Fixed Pay treatment. However, in contrast to that treatment, each participant was paid based on his/her relative performance ranking exactly as in the Private/Fixed Pay treatment. For the two rank-pay treatments, average pay was set at ¥6.00 per round, which was equal to the fixed level of pay in the no-rank-pay treatments.

Table 1 illustrates the payoffs based on rank used in the 20-person employee sessions. The level of pay related to each rank was announced to the participants prior to the commencement of any rank-pay round. A similar payoff structure was used for the 30-person student sessions.[2] We controlled for the mean pay level across all treatments with an average of ¥48.00 from the experimental task component of the study regardless of treatment. This amount appeared salient and meaningful for both types of participants.

Insert Table 1 about here.

Research Participants and Sites

We held seventeen sessions with full-time employees in Zhejiang, a province in southeast China. In a companion study, we carried out six sessions with university students at a medium-

[2] The reason for the different session sizes is discussed in the next section.

sized university in Hangzhou, the capital of Zhejiang province. Table 2 summarizes the distribution of the sessions. As indicated in the table, we ran sessions with three different manufacturing companies (denoted X, Y, and Z for privacy reasons). We ran at least one session of each treatment at each company, and held two of each rank-pay session at Company X. Thus, we have a total of three sessions each for treatments C, Private/Fixed Pay and Public/Fixed Pay and four sessions each for treatments Private/Rank Pay and Public/Rank Pay. There were 20 employee participants in each session for a total of 340 participants. Employee participants were recruited with the assistance of managers in the case of two companies and union officials at another. We were assured that employees knew the other participants in their session well, and this certainly appeared to be the case. With the university student participant pool, we conducted two sessions each for treatments C, Private/Fixed Pay and Public/Fixed Pay, and one session each for the rank-pay treatments, Private/Rank Pay and Public/Rank Pay. There were 40 students in the two control sessions combined and 30 in each of the experimental treatment sessions for a total of 220 student participants. The students in each of the experimental treatment sessions all came from one class, and thus knew each other well. Indeed, the reason that we increased the size of the student sessions from the 20 participants used in the case of employees to 30 participants was to allow a whole class to be involved in each session.[3] Public ranking is likely to matter more within a natural social group such as people working together in a small/medium-sized company or studying together in a Chinese university class.[4]

Insert Table 2 about here.

Table 3 provides a summary of the demographic characteristics of the participants. The mean age of the employee participants was about 28. 61% of them were female, and 36% were married. Their mean monthly income was about ¥1700 (approximately $246.50 US), the mean

[3] The different size of the smaller worker groups versus the larger student groups gives a larger incentive to move up one rank (30 cents versus 20 cents) to the worker groups relative to the student groups. However, as will be reported in the Results section, the effects for workers turn out to be weaker than for students despite the workers having stronger incentives.
[4] In China, classes often take virtually all of their courses together for four years of university education. They also often live together in the same dormitories. This was the case at the university site of our study. Within the separate male and female dorms, four students from the same class shared one room. The rooms of other students from the same class were adjacent.

education level was 2.66 (between junior and senior high school). For the students, the average age was about 20, 56% were female, and all of them had achieved an education level of five (university undergraduate). None were married.

Insert Table 3 about here.

Experimental Procedures

The experimental sessions were held in company conference rooms for employees and classrooms for students. Upon arrival, participants were seated apart from each other and no communication was permitted. The experimental instructions were read aloud to the participants while they followed along on their own printed copies. Each participant was provided with a prepared workbook. For each round, the first page in the workbook explained which compensation scheme and information-feedback condition would apply to the upcoming round. Participants were not permitted to look ahead to future pages or to go back to previous pages. They were only allowed to tear off one page and look at the next when instructed to do so by the experimenter. After each round, each participant's workbook page was collected by the experimenters and taken to another room where the number of correct answers was calculated. In the four non-control treatments, according to the treatment they were in, participants received some or all of the following information after each of the middle four rounds prior to the beginning of the subsequent round: feedback on their ranking information privately or publicly, and their earnings.

After participants completed the experimental task, they filled out a questionnaire in which they responded to demographic questions such as age, gender, marital status and monthly income (the last two items only applied to employee sessions). Besides collecting demographic data, another primary purpose of the questionnaire was to elicit risk preferences.[5] We adopted a risk-aversion elicitation instrument based on Binswanger (1980) and Eckel and Grossman (2008).

[5] We elicited risk preferences after the completion of the experimental task in order to avoid the possibility of biasing the behavioral decisions by priming participants to focus on risk. In our judgment, this was a more important consideration than the possibility that administering the risk attitude measure after completion of the experimental task might lead to an impact of task performance or beliefs about the purpose of the experiment on the subsequent elicitation. Nonetheless, we must acknowledge the latter possibility.

Participants were shown ten circles, depicted in Figure 1. Each circle includes two payoffs, and each payoff occurs with a 50/50 chance. These ten circles are organized in a clock-wise fashion with the top circle containing two identical numbers, representing a certain, risk-free payoff, while the subsequent eight circles represent lotteries that increase in both expected payoff and variance (risk). The last two circles contain lotteries with identical expected payoffs. However, the last circle has a much higher variance to permit identification of participants who may not have risk-averse preferences. Participants were asked to indicate which one of the ten lotteries they would prefer to play for actual monetary earnings. Holt and Laury (2002) found that the amount of money at stake affected risk preferences. In particular, larger stakes were associated with a higher level of risk aversion. We therefore adjusted the stakes in this measure to correspond to the amounts at stake in each round of the real-effort task.

Insert Figure 1 about here.

This risk-attitude elicitation instrument is advantageous for field use for at least two reasons. First, 50/50 gambles are easy to understand and expected payoffs are simple to calculate. Second, the measure is visually presented in a manner that focuses the attention of participants on the fact that the increase in expected earnings is associated with an increase in risk. We thus find this instrument appropriate to use, especially because many of the employee participants had only junior high-school education or less. It has previously been used successfully among Peruvian farmers with limited education (Engle-Warnick et al., 2009, 2011). After participants made and submitted their choices, the experimenter asked a volunteer participant to flip a coin in front of everyone to determine the payoff for each participant according to the circle and consequent gamble he or she selected from the available choices. The earnings of this task were added to the total session earnings for each participant.

At the end of the session, players were taken individually to another room, where they were paid privately in cash. Each session lasted just over an hour. On average, participants earned between ¥50-¥60 from the real-effort task and risk-attitude elicitation combined, which exceeded a day's pay for factory employees and was well in excess of the ¥10-¥15 an hour that students could earn at campus part-time jobs.[6]

BEHAVIORAL HYPOTHESES

Performance Ranking Information as an Inherent, Social-Comparative Incentive

Pioneered by Festinger (1954), social-comparison theory postulates that humans have a fundamental drive to evaluate themselves and that an important source of knowledge about oneself is comparison with other people. When combined with the desire to compare, this drive to improve leads the individual to strive toward better relative performance, leading to greater self-esteem (e.g. Maslow, 1943) and enhanced utility (e.g. Benabou & Tirole, 2002) derived from thinking of themselves as good and valued according to social criteria. Recent behavioral agency theories in economics nicely complement Festinger's early work. For example, Kräkel (2008) uses the concept of emotions to explain the significant oversupply of effort in tournament settings compared to predictions based on the standard assumption that agents care only for their own absolute incomes. Similarly, Ellingsen and Johannesson (2007) argue that people not only appreciate monetary rewards but also get utility from what (they believe) others think about them. In a field quasi-experiment, Blanes i Vidal and Nossol (2011) study the direct effect of informing workers about their own position in the productivity and pay distribution. Since workers in their study were paid by piece-rate, the relative performance information communicated to them privately had no monetary or career consequences. Moreover, its introduction was triggered for exogenous reasons and was not part of any wider managerial policy. Their results show that both *ex ante* anticipation and *ex post* revelation significantly and positively impact productivity, providing strong empirical support for Festinger's social-comparison theory. Moreover, Alos-Ferrer, Garcia-Segarra, and Ritschel (2018) demonstrate experimentally that when given a choice between equal and performance-based pay after performing a real-effort task, the latter was significantly more likely to be selected when it was coupled with the revelation of one's relative performance in contrast to egalitarian pay, which hid such information. They attribute this to

[6] Experimental instructions and materials are available upon request.

curiosity about one's relative performance, which permits social comparison. Echoing these findings, psychologists Buunk and Gibbons (2007), in an informative review of social-comparison theory over five decades, highlight the potential use of social comparison as an instrument to induce positive changes in behavior as an area worthy of more investigation. Specifically, the authors suggest that future research explore the motivational effect of creating an environment that induces social comparison to achieve desirable outcomes.

We take up this theme. We begin by assuming that employees receive flat-wage compensation, which means their pay is not contingent on performance. The provision of private performance-ranking information to such employees provides them with the necessary information to compare their own performance with that of others, thus providing a simple environment conducive to social comparison. We draw these boundaries around our discussion in order to isolate the hypothesized effect of privately conveyed ranking information through social comparison from other potential channels of influence and moderation, two of which we will soon consider as essential elements of our study, while others are beyond the scope of the study.[7] For example, some studies (e.g., Azmat & Iriberri, 2016) consider the receipt of performance-ranking information in a piece-rate environment. We do otherwise for two reasons. First, piece rates alone provide significant financial incentives to perform well, leaving less room for ranking information to have an additional effect. Second, and conversely, in a piece-rate environment, ranking information regarding performance is confounded with ranking information regarding pay since with identical piece rates the higher performer earns more money. Thus, using piece rates, it is impossible to separate the potential effects of learning one's performance rank from the effects of learning one's pay rank because the two are identical[8]. Both may lead to social comparison. However, at this stage, our focus is isolating the effect of mere private performance-rank information on performance when such information has no direct financial implications.

---

[7] See Dechenaux et al. (2015) for a wide-ranging survey of the expansive experimental literature on contests.

[8] Note that So et al. (2017) use piece rates alone and piece rates with ranking information to distinguish between the impact of higher absolute pay alone from a potentially additional impact based on learning one's rank relative to another player. A third two-person winner-take-all-tournament treatment allows the identification of a further possible impact when the higher ranked player wins a predetermined amount while the lower ranked player earns nothing.

H1: Compared to no performance rank information, privately conveyed feedback on performance rank with no direct financial implications has a positive effect on performance.

This is not a novel hypothesis. It has for example been corroborated previously in a similar fixed-wage context using students as participants by Hannan et al. (2012), Kuhnen and Tymula (2012), and Tafkov (2013) as discussed in more detail after the presentation of H2. In contrast to earlier research, we extend the domain of the study and seek replication to include not only university students, but also full-time employees to assess the robustness of this behavioral prediction beyond the student to the worker population. Moreover, an examination of this hypothesis is a necessary baseline to study the several other hypotheses with which we are principally concerned in this study.[9]

Public Ranking as a Moderator of the Impact of Performance Rank on Performance

When rank information is provided privately, the effect of social comparison on performance can only function through the individual inclination toward self-evaluation and improvement discussed by Festinger (1954). However, with public ranking, each employee knows that his or her performance rank will be announced publicly to all members of his or her peer group. This opens an additional channel through which rank information may affect performance. Now, in addition to exerting effort in order to experience positive feelings about one's own achievements relative to others, an individual is exposed to the possibility of

[9] Two studies examine a different but related question, namely whether the frequency or level of detail of the feedback affects performance under piece rates and under a winner-take-all tournament. For example, in all three treatments compared by Hannan et al. (2008) participants receive relative performance information about the decile level of their performance when they have finished the experimental task. The authors show that receiving feedback at quarterly intervals on either one's performance decile (called fine) or whether or not one is above or below the 50th percentile (called course) results in improved performance compared to receiving feedback only at the end (called no feedback) under piece rates, but a deterioration in performance under a winner-take-all tournament. Eriksson et al. (2009) also compare three feedback regimes under both piece rates and a winner-take-all tournament in a real-effort laboratory experiment in France. In all three treatments, two subjects are paired and in all of them subjects are informed of their counterpart's score when the task is completed. Thus, all three treatments involve relative performance feedback. The treatments differ in terms of how frequently such feedback is provided while the subjects are working on the task. In one treatment, no feedback is provided while the task is underway (no feedback), in the second it is provided halfway through the task (discrete feedback), and in the third it is provided on a continuous basis (continuous feedback). There is no significant difference in average performance between these treatments under piece rates. There is some evidence of positive peer effects for continuous feedback under the tournament scheme.

establishing high or low status in front of his or her peers, and the pride or embarrassment that results. This provides a potentially second incentive to perform well.

This line of argument is consistent with Frank (1985), who notes that local status, i.e., status within one's reference group, affects behavior. Moreover, he argues that communicating one's ability to others is important in society. However, since ability is unobservable, it must be signaled through behavior. Performance ranking feedback communicated publicly can thus introduce a magnified social-comparative incentive if people are driven to maintain a positive impression among their peers. In other words, public relative performance ranking information represents public recognition and can function as a non-financial reward based on job performance, and act as a social, non-monetary incentive for better performance in addition to the incentive provided by private ranking information alone.

H2: Compared to privately provided feedback on rank with no financial implications, public ranking has a positive moderating effect on the relationship between rank information and performance.

To the best of our knowledge, all previous studies involving the provision of complete ranking information with a fixed-salary baseline have used only students as participants (Charness et al., 2014; Hannan et al., 2012; Kuhnen & Tymula, 2012; Tafkov, 2013). Thus, we maintain that it is of critical importance to consider the extent to which these results consistent with *H1 and H2* generalize to employees. Closely related studies suggest there may be important differences between populations. For example, while studies have demonstrated that both private and public ranking information improve performance under piece rates (Blanes i Vidal & Nossol, 2011; Hannan et al., 2008; Tavkov, 2013), two studies by Barankay (2011, 2012) both find that private ranking information weakens performance among MTurk workers and furniture salespeople respectively in a piece-rate environment. While these studies differ in many details, and are thus not strictly comparable to each other, they nonetheless suggest the need to corroborate the existing results on the effects of ranking information and the moderating effect of public ranking under fixed salaries on non-student populations such as factory employees.

Moving beyond student participants is an important goal of our research, and we would argue that such replication using different populations is key to the advancement of knowledge.

Rank Pay as a Moderator of the Impact of Performance-Rank Feedback on Performance

Rank pay, while providing financial incentives for all, creates a clear economic hierarchy. Consistent with social-comparison theory, Frank (1985) notes that concerns about one's position in the economic hierarchy can affect one's behavior. Clark et al. (2008) survey literature on relative income and happiness, and report empirical evidence of a positive relationship between them. Similarly, Knight et al. (2009) present empirical evidence from villages in rural China, demonstrating a significant relationship between income relative to one's fellow village residents and reported levels of happiness. Moreover, simply earning more even in absolute terms in the experiment clearly affords greater command over resources and hence higher utility according to standard economic theory. Our rank pay treatments allow us to test the additional social comparative effect of relative income linked to ranking, above and beyond mere ranking information with no financial consequences, on performance. This leads to our third behavioral hypothesis:

H3: Compared to fixed pay, rank pay has a positive moderating effect on the relationship between rank information and performance.[10]

Individual Risk Attitude as a Moderating Variable

In considering the decision of how much effort to exert when performance-rank information is to be provided, employees need to take account of the riskiness inherent in their decisions. Since individuals in general have differing levels of risk tolerance, they may make different decisions in an identical situation. Hence, it is important to examine the moderating effect of different attitudes toward risk.

Consider an employee who is very averse to risk. Both Skaperdas and Gan (1995) and Anderson and Freeborn (2010) discuss two intuitions with contrasting implications. The first is

[10] An alternative perspective is provided by the literature on incentive crowding-out effects (e.g., Frey & Jegen, 2001; Gneezy & Rustichini, 2001a, 2001b).

that a more risk-averse person might choose to exert less effort in pursuit of a risky prospect than a person who is less risk-averse. This is because for the more risk-averse person, the expected utility gain from such a risky prospect is less than for somebody who is less risk-averse. Thus, there is a reduced incentive to bear the certain disutility cost of exerting effort. This intuition suggests that higher levels of individual risk-aversion will be associated with less performance improvement from the receipt of information on performance rank than lower levels of risk aversion regardless of whether ranking feedback is private or public, or whether pay is fixed or based on rank. In contrast, the second intuition is that a more risk-averse person might be inclined to exert more effort in order to protect him/herself against the risk of failing to achieve a high rank.

These contradictory intuitions have motivated economic theorists to develop a number of mathematical models, which together demonstrate that whether more risk-averse people will exert more or less effort than less risk-averse people in a contest depends on the detailed rules of the contest and the specific assumptions made about the shape of individual utility functions (e.g., Cornes & Hartley, 2003; Hillman & Katz, 1984; Hillman & Samet, 1987; Konrad & Schlesinger, 1997; Millner & Pratt, 1991; Skaperdas & Gan, 1995; Treich, 2010). Collectively this literature suggests that while in general greater risk aversion can lead to either less or more effort, in many contests the most commonly used utility functions are often consistent with the intuition that associates risk aversion with the exertion of less effort. For example, Treich (2010) demonstrates that risk aversion reduces effort relative to risk neutrality if utility functions exhibit prudence, i.e. a positive third derivative. Prudence is implied by absolute risk-aversion that decreases with wealth and this is a property of many common utility functions including the well-known Constant Relative Risk Aversion function.

However, since economic theory allows for either intuition to dominate, which one actually does boils down to an empirical question. We know of three laboratory experiments that address this issue in a variety of different contest formats (Anderson & Freeborn, 2010; Sheremeta, 2011; Sheremeta, Masters & Cason, 2012). They all find a significant inverse relationship between risk aversion and effort, suggesting the dominance of the first intuition. However, effort in these studies is not the exertion of real effort but rather effort represented by monetary bids that are positively correlated with the probability of winning a prize. Moreover, none of these contests involve ranking feedback or rank pay. A fourth experiment was conducted in an educational context where grades on a portion of an exam were based on whether one ranked first or second relative to a randomly assigned counterpart (De Paola et al., 2016). The authors find that more risk-averse participants perform worse in response to a larger spread between the winning and losing prizes with a constant expected value presumably because they exert less effort, while the performance of less risk-averse participants is not affected. Our reading of the theory buttressed by the existing experimental evidence suggests that in the context of the social comparison that arises from the receipt of performance-rank information, individuals with higher levels of risk aversion will exert less effort when exposed to risk than those with less aversion to risk. This attenuates the positive impact of the social comparison induced by the performance-rank information on performance.

H4: Individual levels of risk aversion have a negative moderating effect, attenuating the positive relationship between rank information and performance.

H4 does not differentiate between the potential quantitative impact of the moderating effect of risk aversion when the utility of being ranked high is psychologically based on shame and pride alone versus when it is also financially based on rank-linked pay. Similarly, it does not differentiate between the predicted impact in the case of privately versus publicly provided information. We might expect the moderating impact of risk aversion to be greater in those situations where social comparison is predicted to have the strongest effects on utility and hence behavior, since in those situations the riskiness associated with any given random component of rank determination will also have a bigger impact on utility. This would imply that the moderating effect of risk aversion would be larger when psychological and financial factors are combined under rank pay than when an employee faces risk from psychological factors alone under fixed pay. Similarly, it would suggest that the moderating effect of risk aversion would be larger with public ranking than with the private transmission of rank information. We examine these intuitions empirically.

RESULTS

Aggregate Results

A preliminary look at the data revealed that the results for each treatment appeared to be quite similar at each of the three company sites. To test for possible systematic differences among the three sites we conducted a series of joint Chi-squared tests utilizing our two dependent variables: Performance Improvement (*Per_Imp*), defined as the difference between the average number of problems solved in the middle four rounds and the average number solved in the first two rounds and Further Improvement in the last two rounds (*Fur_Imp*), defined as the difference between the average number of problems solved in the final two rounds and the average number of problems solved in the middle four rounds. For each of these variables, we conducted five tests, one for each treatment. The null hypothesis in each case was of equality among all three company sites. This null hypothesis was not rejected for either variable in any of the treatments. Thus, we pooled the employee data for the subsequent analysis.[11]

Table 4 presents average per-round performance over all eight rounds, as well as average *Per_Imp* and *Fur_Imp* for both students and employees by treatment. The final row of the table presents our measure of risk-aversion for the sample from each population. The responses were coded as integers, with the highest number 9 assigned to the safest, most risk-averse choice at the top of the instrument illustrated in Figure 1. Each subsequent circle was indexed by subtracting one, moving clockwise around the larger circle to the last small circle, denoting a risk-loving or risk-neutral choice, coded as 0. Thus, a larger integer indicates a more risk-averse response to the elicitation instrument. The table reveals several broad stylized facts regarding the experimental results. First, even in the absence of rank information, performance in the control condition was significantly positive. Second, average student performance was higher than average employee

[11] Although not necessary to justify the pooling of the data from the three different companies, we also failed to reject the null hypothesis that the average level of risk aversion was identical across the three sites. Moreover, we could not reject the null hypothesis that the average per-round performance level was identical for each individual treatment across the three sites.

performance in every treatment (broadly, students solved approximately 12 problems, while employee performance was in the neighborhood of seven). Third, across both the control and all experimental treatments both populations exhibited increases in performance during the middle four rounds compared to the first two rounds. Fourth, employees made lottery decisions that were on average one full "circle" more risk-averse than students.[12]

Insert Table 4 about here.

Treatment Effects and Behavioral Results

We report treatment effects using the focal performance measure of performance improvement (*Per_Imp*) between the first two and the middle four rounds. In each case, there is one observation for each participant. The tests are performed in a regression framework. We used a maximum-likelihood random effects model to control for possible unobserved differences between sessions. For both employees and students, a Likelihood Ratio test with the null hypothesis that there were no such differences was equal to 0.000, providing no evidence of any such session-related effects. We consider *Per_Imp* to be the most appropriate measure for treatment effects because it controls for individual differences in baseline ability and motivation that may affect performance throughout the session. We perform separate regressions for students and employees. Dummy variables are used to represent potential differences between each experimental treatment and the control treatment as follows: the *private/fixed pay* dummy is one for the Private/Fixed Pay treatment and zero otherwise, the *public/fixed pay* dummy is one for the Public/Fixed Pay treatment and zero otherwise, the *private/rank pay* dummy is one for the Private/Rank Pay treatment and zero otherwise, and the *public/rank pay* dummy is one for the Public/Rank Pay treatment and zero otherwise. *Risk Aversion* (*RA*) is a variable representing individual levels of risk-aversion. We also employ interaction variables between RA and each of

[12] For both students and employees, pairwise Kolgomorov-Smirnoff tests could not reject the null hypothesis that elicited risk-aversion measures were drawn from the same distribution for any two treatments being compared.

the treatment dummies. The regression results are reported in Table 5, Panel A separately for student and employee participants.[13]

Despite the unidirectional predictions of our hypotheses, we conservatively employ two-tailed tests throughout our analysis. Post-estimation Chi-squared tests are reported in Table 5, Panel B in order to compare the experimental treatments with each other. Results are numbered to link them to the behavioral hypotheses in the previous section.

Insert Table 5 about here.

H1 Tests: As predicted by Hypothesis 1, compared to the control treatment (with no rank information and fixed pay), providing rank information privately had a significantly positive effect on performance for both students and employees.

The coefficient on *private/fixed pay* is positive and significant for both students (1.79, p <0.01) and employees (0.76, p=0.02), thus corroborating the predicted positive effect of performance ranking information on performance improvement (*Per_Imp*) in the private/fixed pay treatment compared to the control.

Moreover, for students, compared to the control treatment (with no rank information and fixed pay), providing rank information also has a significantly positive effect on performance in the other three experimental treatments. The coefficients on the *public/fixed pay* (1.70, p<0.01), *private/fixed pay* (1.08, p=0.04), and *public/rank pay* (2.49, p<0.01) dummies are all positive and significant for students. However, for employees, providing rank information has a marginally significantly positive effect in the public/fixed pay treatment, a significantly positive effect in the private/rank pay treatment, and a positive but insignificant effect in the public/rank pay treatment.

---

[13] Gender was initially entered as a control variable. It was never significant. None of the other demographic variables (age, marital status or income of employees) were significant either, whether examined individually or jointly. Moreover, the AIC and BIC statistics, which are attempts to trade off greater parsimony with additional explanatory power are both smaller (which is better) without the demographics than with the demographics. Therefore, we dropped the demographic controls from the analysis. It must be noted that while all but three of the student participants responded to all of the demographic questions, in our population of 340 factory workers, only 218 provided all the demographic information. This implies that the tests we allude to above were by necessity done only using the subset of subjects who provided the demographic information. Therefore, strictly speaking our statistical argument for dropping the demographic variables only applies to this subset of subjects. In order to use all the data, we need to assume that the demographic controls were also insignificant for the subjects who failed to provide their demographic information to us. In our judgment, this assumption, while untestable, is reasonable, and the cost of not making it is large. Hence we make this assumption and use the data from all the subjects in the paper.

These corresponding coefficients are 0.53 (*p*=0.10), 0.62 (*p*=0.04), and 0.40 (*p*=0.13) respectively. It is puzzling that the impact of ranking feedback on performance was significantly greater in the Public/rank pay treatment than in the control for students, but though positive, not significant for employees. We discuss this further immediately after our discussion of the H3 tests below.

H2 Tests: Contrary to Hypothesis 2, which predicted that compared to private ranking, public ranking would have a positive moderating effect on the relationship between rank information and performance, there was no such significant moderating effect under fixed pay for either students or employees.

Chi-squared test 1 compares the public/fixed pay treatment to the private/fixed pay treatment. It shows no significant effects either for students (*p*=0.81) or for employees (*p*=0.49). Hence, we conclude that these two treatments are not significantly different from each other.

Moreover, for students, there is a significantly positive moderating effect of public ranking on the relationship between ranking feedback and performance under rank pay. However, this is not the case for employees, for whom there is no significant moderating effect under rank pay. Chi-squared test 4 compares the public/rank pay treatment to the private/fixed pay treatment. It yields a significantly positive effect for students (p=0.01), but not for employees (p=0.37).

H3 Tests: Contrary to Hypothesis 3, which predicted that compared to fixed pay, rank pay would have a positive moderating effect on the relationship between rank information and performance, there was no such significant moderating effect in the private ranking treatments for either students or employees. In the public ranking treatments, however, there is a marginally significant moderating effect of rank pay on the relationship between ranking feedback and performance for students, but not for employees.

Chi-squared test 2 compares the private/rank pay treatment to the private/fixed pay treatment. It fails to reject the null hypothesis of no moderating effect for both students (*p*=0.13 and employees (*p*=0.64). Chi-squared test 3 compares the public/rank pay treatment to the public/fixed pay treatment. It yields a marginally significant positive effect for students (*p*=0.09), but not for employees (*p*=0.37). Together the results of these tests reflect an especially strong effect of receiving rank information for students in the public/rank pay treatment where publicly announced rank information implies publicly announced salary differentials commensurate with performance. It would seem that students are especially motivated when they know everyone else will be able to see not only their performance ranking, but also their pay ranking and indeed their pay level since each rank is linked with a pre-announced level of pay. Thus higher pay appears to be associated with higher status within the class for students. No such effect is apparent for employees. This turns out to be the only significant effect of the proposed organizational moderators in our data.

Of course, failing to find an effect does not automatically mean that the null hypotheses is true. Therefore, we undertake a power analysis to assess the likelihood of finding such effects given that they exist, and how this likelihood relates to the true effect size. Assuming a significance level of 5%, we calculated the minimum detectable effect size for power levels of 0.7, 0.8, and 0.9. The greater number of observations for employees (n=340) compared to students (n=220) implies that the minimum detectable effect size is lower for employees than for students at each power level. Effect size is specified as Cohen's f^2, which is defined as $(R_c^2 - R_r^2)/(1 - R_c^2)$, where R_c^2 is the explained sum of squares over the total sum of squares for the complete model and R_r^2 is analogous for the reduced model excluding the one explanatory variable that is the object of the power analysis (Cohen, 1988). Intuitively, this definition of effect size represents the change in the explained sum of squares resulting from the addition of the variable under analysis as a proportion of the unexplained sum of squares from the complete model inclusive of the variable being analyzed.

The results of the power analysis indicate that for students the minimum detectable effect size ranges from 0.028 with power 0.7 to 0.036 with power 0.8 to 0.048 with power 0.9 for a two-tailed test. For employees, the comparable numbers are 0.018 with power 0.7 to 0.023 with power 0.8 to 0.031 with power 0.9 for a two-tailed test. For students, $R_c^2 = 0.19$ while for employees $R_c^2 = 0.24$.[14] While our tests would lack power against explanatory variables with a very small effect size, they have considerable power to detect substantive effects on the proportion of variance explained by the proposed moderators. With the exception of students under the public/rank pay treatment, there is no evidence of such effects.

H4 Tests: As predicted by H4, individual levels of risk aversion had a negative moderating effect, attenuating the positive relationship between rank information and performance.

For individual levels of risk aversion (*RA*), the moderating effects turn out to be much more important. H4 suggests negative effects on *Per_Imp* in all of the experimental treatments that involve the receipt of rank information. This excludes the control treatment in which no rank information is given to participants. The coefficient on *RA* tests the null hypothesis for the control treatment and finds that as predicted it cannot be rejected either for students (p=0.13) or for employees (p=0.11). In contrast, the coefficients on all the interactions between *RA* and the four treatments are negative and significant albeit in one case marginally so for both students (p<0.01 for its interactions with *private/fixed pay*, *public/fixed pay*, and *public/rank pay* and p=0.07 for its interaction with *private/rank pay*) and employees (p=0.04, 0.05, 0.01 and 0.03 for its interaction with *private/fixed pay*, *public/fixed pay*, *private/rank pay*, and *public/rank pay* respectively). These results suggest that participants who are more risk-averse exhibit less of a performance improvement in response to rank information than those who are less risk-averse compared to the control where there is no significant difference in performance improvement based on risk attitude. This is true for all of the experimental treatments. Thus, *RA* is an attenuating moderator on the effect of rank information on *Per_Imp* regardless of the rank transparency or rank pay condition. Chi-squared test 5 at the bottom of Table 5 shows that there is no significant difference in the size of this effect for either students (p=0.43) or employees (p=0.89).

[14] Strictly speaking, our mle random-effects estimation does not yield an R^2 statistic. However, a Likelihood Ratio test with the null hypothesis that there were no session-related random effects was equal to 0.000 with a resulting p-value of 1 for both students and employees. Re-estimating with OLS yields identical coefficients and almost identical standard errors. We use the R^2 statistics from this OLS estimation in the power calculations.

The significance of *RA* as an attenuating moderator is especially important because attitudes toward risk may differ across populations. The effectiveness of rank information at improving performance will in general depend on such attitudes. The two populations upon which we drew for our study highlight this issue. The mean level of risk aversion for students was 2.88, while for employees it was significantly higher at 3.87 ($p<0.01$). The median levels of risk-aversion for students and employees were also significantly different at 1 and 4 respectively ($p=0.01$), while the modal choice was 0 for both students and employees.[15] To illustrate the importance of this difference in the distribution of risk attitudes between populations, we redo the previous analysis, but drop the *RA* moderator so that the estimates will reflect the population average level of *RA* in each case. The results are reported in Table 5 alongside the full model already discussed. For students, the effect of rank information stays intact, while for employees it disappears.

The implications of this analysis are critical. If we ignore risk preferences, and examine only the average response, students seem to improve their performance significantly in reaction to rank information, while employees do not. However, when we control for the attenuating effect of risk aversion on the effect of rank information on *Per_Imp*, we find that such information does result in higher *Per_Imp* for both students and employees compared to the control when aversion to risk is sufficiently low. Correspondingly, rank information fails to result in higher *Per_Imp* for either students or employees when risk aversion is sufficiently high. Thus, risk preference is a major factor in determining the effectiveness of social comparison primed by rank information at improving performance.

Despite the difference in the median and mean levels of risk aversion, within both groups there are many individuals whose risk-aversion levels are low. For example, for both samples, the mode for risk aversion is at the lowest possible level of zero, this being true for 31% of students and 23% of employees. Such people will be particularly prone to improve their performance

[15] The reasons for different attitudes toward risk between our two populations are beyond the scope of this paper. However, we could find no significant relationships within the student or worker populations between elicited measures of risk aversion and any of the demographic variables.

when they receive the ranking information needed to make social comparisons whether they are students or employees.[16]

Supplementary Analysis: Removing the Social-Comparison Incentives

Recall that in the last two rounds of all treatments, all feedback on ranking was removed. As in the first two rounds, participants were paid a fixed amount that did not depend on performance and rank information was no longer provided. We are interested in the question of whether and how social-comparison incentives based on ranking information in place at one time might affect subsequent performance once they are removed (e.g., Neckermann et al., 2014). If the social-comparison incentives primarily affected the participants by motivating them simply to work harder or at a faster pace, one might expect that participants would revert to lower levels of effort once such incentives were withdrawn. However, if the social-comparison incentives caused participants to learn how to perform the task more effectively, the performance levels induced by such incentives might outlive their presence. The *Fur_Imp* numbers are reported in Table 4. In the control treatment, performance falls for both students and employees, though the fall is not significant for the students (p=0.44) and only marginally so for the employees (p=0.09). For students (employees), *FI* was positive for three of the four (all four) experimental treatments. In both instances, for the two treatments in which rank was not linked to pay, private/fixed pay and public/fixed pay, the results were statistically significant (for students p=0.02 and p=0.05 respectively and for employees p=0.01 and p<0.01 respectively).[17]

We then employ *Fur_Imp* as the dependent variable in two regression models estimated

[16] We also examined whether participants were reacting to the announcement that a new information/compensation environment would apply made at the beginning of period 3 as assumed throughout this section and/or to the receipt of specific ranking information (and consequent earnings in the rank-pay treatments) that first occurs at the end of period 3. A recent innovative study by Gill et al. (2017) exploits data on ties broken randomly in allocating ranks to identify a U-shaped rank response function such that subjects exhibit 'first-place loving' and 'last-place loathing', working hardest after being ranked either first or last. We were unable to exploit their identification technique because we did not assign ties to ranks randomly. With the available statistical techniques, we could find no systematic reaction to the receipt of a particular rank or payment based on it. In particular, there was no difference by ranking quartile comparable to Gill et al. (2017) or otherwise. Moreover, the variance of performance levels over all the ranked participants did not change significantly over time as would occur if lower ranked performers improved more while higher ranked performers improved less. Rather, participants reacted similarly from period 3 onward to the announcement that a new information/compensation condition would henceforth prevail. These tests are not reported in detail to save space, but are available from the authors upon request.

[17] For the private/rank pay treatment, the p-values were 0.03 for students and 0.07 for employees. For the public/rank pay treatment, the coefficient was negative with a p-value of 0.76 for students and positive with a p-value of 0.06 for employees.

using maximum likelihood to include a random effect for session. The results are reported in Table 6. Model 1 includes all the independent variables that we employed in the performance-improvement model 1 regression reported in Table 5. In addition, it includes each individual's performance improvement that occurred previously moving from the first two to the middle four rounds (*Per_Imp*) as an additional explanatory variable. The idea is that further improvement might be mitigated by earlier improvement since there is presumably a limit on how much each person is capable of improving. In this model, the only explanatory variable that is significant is *Per_Imp*, which is negative and marginally significant for students (p=0.08), and negative and significant for employees (p<0.01).

Insert Table 6 about here.

Post-estimation Chi-Squared tests indicate that there are no significant differences between the four rank-feedback treatments in Model 1. If we aggregate all four treatments into one dummy variable called FEED and drop the insignificant risk-aversion variable and its insignificant interactions, we obtain the stripped-down Model 2, a regression model with two independent variables: the dummy variable indicating the availability of feedback on rank (*FEED)* and the improvement that occurred previously moving from the first two to the middle four rounds (*Per_Imp*). For students, the coefficient on *FEED* is 0.72 (p=0.05), while the coefficient on *Per_Imp* is −0.17 (p=0.04). For employees, the corresponding numbers are 0.65 (p<0.01) and −0.23 (p<0.01). Our interpretation is that in a difference-in-difference analysis compared to the control treatment, performance continued to improve from the middle four to the last two rounds after the rank-feedback incentives were removed in the four experimental treatments for both students and employees regardless of the proposed moderating variables. However, the magnitude of such continued improvement was mitigated by the size of the improvement that had already occurred from the first two to the middle four rounds.[18] We conjecture this represents a ceiling effect on each person's ability to perform well.

[18]An interaction between *FEED* and *Per_Imp* was not significant either for students or employees. The detailed results of these regressions are not reported here in order to save space, but are available from the authors upon request.

CONCLUSION

Contributions

We examine the relationship between the provision of performance-ranking information and task performance as well as various contingencies that might affect that relationship through the suggestive lens of social-comparison theory. Our focus is on three potential moderating variables: private versus public ranking (whether ranking information was released privately to each individual employee or announced publicly to all); fixed versus rank pay (whether ranking was financially consequential or not); and individual risk attitude. We make several important contributions.

First, we extend the empirically demonstrated boundaries of social-comparison theory in the context of performance-rank feedback beyond the university laboratory to the factory floor, while simultaneously establishing several important differences in the roles played by our proposed organizational moderators among the populations in these two environments. In particular, while public ranking and rank pay together intensify the relationship between feedback and performance when they are both present among students, they have no effect at all among the factory employees in our study. For the students in our study, this suggests that the additional push to performance predicted by status concerns (Frank, 1985) comes into play only when coupled with a link to pay.

Second, our results suggest that the magnitude of the response to performance ranking depends on individual heterogeneity, and in particular on individual attitudes toward risk as an important moderator of that response. It thus extends and contributes to the continuing theoretical-empirical interplay concerning the relationship between risk attitudes and effort exertion in contests. To our knowledge, our study represents the first to test this using a real-effort task in a work context. It also gives strong support to an inverse relationship between risk aversion and performance, which reflects an inverse relationship between risk aversion and effort as hypothesized.

Finally, in terms of managerial implications, we demonstrate that individual attitudes toward risk can significantly impact how effective a rank-pay incentive may be at improving productivity. We show that such an impact may occur not only with the introduction of financial incentives, but also with the use of performance-ranking information, whether or not such information is related to pay. In particular, when given feedback on their performance rank, more risk-averse employees are likely to be less responsive, unresponsive, or even respond in the wrong direction. Since the stakes are bound to be higher in the workplace than in our experiment, and risk aversion is likely to be greater at higher stake levels (Holt & Laury, 2002), this phenomenon could well be even more pronounced in the workplace than in our study. Thus, different kinds of pay schemes may suit different kinds of employees, and risk attitudes may be a critical factor in determining the best employee-compensation fit.

Limitations and Issues for Further Study

As in every study, we had to make design choices to focus on issues of primary concern at the cost of not examining all related issues. Moreover, we were constrained by the number of factory employees available and had to ensure that the demands of the experiment were not beyond their cognitive capacity.

First, our design, like many influential papers (Azmat & Iriberri, 2016; Belogolovsky & Bamberger, 2014; Bradler et al. 2015; Cadsby et al., 2007; Charness et al. 2014; Dickinson & Villeval, 2008; Dutcher et al. 2015; Eriksson et al. 2009; Falk & Ichino, 2006; Freeman & Gelber, 2010; Gill et al. 2017; Hannan et al. 2008, 2012; Kosfeld & Neckermann, 2011; Kuhnen & Tymula, 2012; Tafkov, 2013), did not include an alternative leisure activity. Corgnet et al. (2015) and Engel (2010) show that having an alternative activity can lead to differing results and in particular stronger incentive effects than when such an alternative activity is absent. One reason we did not utilize an alternative leisure activity was for comparability with this mainstream literature. A second reason is that although many jobs allow for shirking via leisure activities such as browsing the Internet, many others do not have such possibilities available. In the factories that were the setting for our research project, subjects worked on a machine doing a task and did not have access to a computer. In our view, an experimental setting without an alternative leisure-based activity better represented this setting. A third reason is that it would have been difficult to find a leisure activity of similar value to our worker and student subjects, especially in the absence of a computer interface at the factories. There is however no doubt that examining how the effects of incentives based on social comparison might differ when an alternative leisure activity is available is an important issue for further study.

Second, some other experimental design features may also have influenced our results. We adopted the number-addition task, a standard real-effort task in the literature, in our experiment. It is worth noting that solving arithmetic questions quickly and accurately is likely to be a skill especially valued by students at the "industry and commerce" university they attended. Hence they may care more about what others can observe regarding their skill at this task and the money they can earn based on that skill than the less educated and educationally oriented workers. For workers, moreover, earnings from their actual job were likely more important than earnings based on rank in the experiment as an indicator of status. It appears that differences in earnings resulting from rank pay in the experiment meant little more than performance ranks in such a context.

Third, we used a risk-elicitation method that has proven to be easily understood among less educated populations (Binswanger, 1980; Eckel & Grossman, 2008; Engle-Warnick et al., 2009, 2011). The disadvantage, however, is that while distinguishing between differing levels of risk aversion, it does not distinguish between differing levels of risk loving since all subjects who are risk loving should select the same option. Risk-neutral subjects would be indifferent between two options, one of which is also the risk-loving option. In our opinion, this lack of precision for subjects who are not risk-averse was less important than having a risk-elicitation instrument with a proven track record among less educated subjects in the field. For similar reasons, we did not measure or analyze the effect of ambiguity nor did we examine how expectations about one's rank might affect the exertion of effort. The latter would likely be a bigger issue in a winner-take-all tournament because less confident people are more likely to give up on winning the grand prize and reduce effort accordingly. Our subjects can always hope to increase their rank whether from a low or a high baseline expectation. These are all issues for further study.

Acknowledgement: We gratefully acknowledge the Social Sciences and Humanities Research Council of Canada for funding this research (Grant #410-2010-1174).

References:

Alos-Ferrer, C., Garcia-Segarra, J. & Ritschel, A. 2018. Performance curiosity. ***Journal of Economic Psychology***, 64, 1–17.

Anderson, L., & Freeborn, B. 2010. Varying the intensity of competition in a multiple prize rent seeking experiment. ***Public Choice***, 143, 237–254.

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

Barankay, I., 2011. Rankings and social tournaments: Evidence from a crowd-sourcing experiment. Unpublished manuscript.

Barankay, I. 2012. Rank Incentives: Evidence from a Randomized Workplace Experiment. Unpublished manuscript.

Belogolovsky, E., & Bamberger, P. A. 2014. Signaling in secret: Pay for performance and the incentive and sorting effects of pay secrecy. ***Academy of Management Journal***, 57, 1706–1733.

Benabou, R., & Tirole, J. 2002. Self-confidence and personal motivation. ***Quarterly Journal of Economics***, 117: 871–915.

Binswanger, H. P. 1980. Attitudes toward risk: Experimental measurement in rural India. American. ***Journal of Agricultural Economics***, 62, 395–407.

Blanes i Vidal, J., & Nossol, M. 2011. Tournament without prizes: Evidence from personnel records. ***Management Science***, 51, 1721–1736.

Bradler, C., Dur, R., Neckermann, S., Non, A. 2015. Employee recognition and performance: A field experiment. ***Management Science***, 62, 3085–3099.

Buunk, A. P. & Gibbons, F. X. 2007. Social comparison: The end of a theory and the emergence of a field. ***Organizational Behavior and Human Decision Processes***, 102, 3–21.

Cadsby, C. B., Song, F., & Tapon, F. 2007. Sorting and incentive effects of pay for performance: An experimental investigation. ***Academy of Management Journal***, 50, 387–405.

Cadsby, C. B., Servatka, M., & Song, F. 2013. How competitive are female professionals? A tale of identity conflict. ***Journal of Economic Behavior and Organization***. 92, 284–303.

Cadsby, C. B., Song, F., & Tapon, F. 2016. The Impact of risk aversion and stress on the incentive effect of performance pay. In S. Goerg and J. Hammon (Eds.), ***Research in Experimental Economics Vol.19: Experiments in Organizational Economics***, 189–227.

Charness, G., Masclet, D., & Villeval, M. C. 2014. The dark side of competition for status. ***Management Science***, 60, 38–55.

Clark, A., Frijters, P., & Shields, M. 2008. Relative income, happiness, and utility: An explanation for the Easterlin paradox. ***Journal of Economic Literature***, 46, 95–144.

Cohen, J. 1988. ***Statistical Power Analysis for the Behavioral Sciences***. (2nd Edition). Mahwah, NJ: Erlbaum.

Corgnet, B., Hernán-González, R., & Schniter, E. 2015. Why real leisure really matters: incentive effects on real effort in the laboratory. ***Experimental Economics,*** 18, 284-301.

Cornes, R., & Hartley, R. 2003. Risk aversion, heterogeneity, and contests. ***Public Choice***, 117,

1–25.
Dechenaux, E., Kovenock, D., & Sheremeta, R. 2015. A survey of experimental research on contests, all-pay auctions and tournaments. ***Experimental Economics***, 18, 609–669.
De Paola, M., Gioia, F., & Scoppa, V. 2016. The adverse consequences of tournaments: Evidence from a field experiment. *University of Calabria Working Paper 07-2016*.
Dickinson, D., & Villeval, M. C. 2008. Does monitoring decrease work effort: The complementarity between agency and crowding-out theories. ***Games and Economic Behavior***, 63, 56-76.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G.G. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. ***Journal of the European Economic Association***, 9, 522–550.
Dutcher, E. G., Balafoutas, L., Lindner, F., Ryvkin, D., & Sutter, M. 2015. Strive to be first or avoid being last: An experiment on relative performance incentives. ***Games and Economics Behavior***, 94, 39–56.
Eckel, C. C., & Grossman, P. J. 2008. Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. ***Journal of Economic Behavior and Organization***, 68, 1–17.
Engel, R. 2010. Why work when you can shirk? Worker productivity in an experimental setting. ***Journal of Applied Business and Economics***, 11, 104-119.
Ellingsen, T., & Johannesson, M. 2007. Paying Respect. ***Journal of Economic Perspectives, 21***, 135-149.
Engle-Warnick, J., Escobal, J., & Laszlo, S. 2009. How do additional alternatives affect individual choice under uncertainty? ***Canadian Journal of Economics***, 42, 113–140.
Engle-Warnick, J., Escobal, J., & Laszlo, S. 2011. Ambiguity aversion and portfolio choice in small-scale Peruvian farming. ***B.E. Journal of Economic Analysis and Policy***, 11, Article 68.
Eriksson, T., Poulsen, A., & Villeval, M. C. 2009. Feedback and incentives: Experimental evidence. ***Labour Economics***, 16, 679–688.
Falk, A., & Ichino, A. 2006. Clean evidence on peer effects. ***Journal of Labor Economics,*** 24, 39–57.
Festinger, L. 1954. A theory of social comparison processes. ***Human Relations***, 7, 117–140.
Frank, R. 1985. ***Choosing the Right Pond.*** New York: Oxford University Press.
Freeman, R., & Gelber, A. 2010. Prize structure and information in tournaments: Experimental Evidence. ***American Economic Journal: Applied Economics***, 2, 149–164.
Frey, B., & Jegen, R. 2001. Motivation Crowding Theory. ***Journal of Economic Surveys***, 15, 589–611.
Gill, D., Kissova, Z., Lee, J., & Prowse, V. 2017. First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision. Working Paper.
Gneezy, U., & Rustichini, A. 2001a. Pay enough or don't pay at all. ***Quarterly Journal of Economics***, 115, 791–810.
Gneezy, U., & Rustichini, A. 2001b. A fine is a price. ***Journal of Legal Studies***, 29, 1–18.
Hannan, R. L., Krishnan, R. & Newman, A.H., 2008. The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. ***The Accounting Review***, 83, 893–913.
Hannan, R. L., McPhee, G. P., Newman, A. H. and Tafkov, I. D., 2012. The effect of relative performance information on performance and effort allocation in a multi-task environment. ***The Accounting Review***, 88, 553–575.

Harrison, G. W., & List, J. A. 2004. Field experiments. ***Journal of Economic Literature***, 42, 1009–55.
Hillman, A.L., & Katz, E. 1984. Risk-averse rent seekers and the social cost of monopoly power. ***Economic Journal***, 94, 104–110.
Hillman, A.L., & Samet, D. 1987. Dissipation of contested rents by a small number of contenders. ***Public Choice***, 54, 63–82.
Holt, C., & Laury, S. 2002. Risk Aversion and Incentive Effects. ***American Economic Review***, 92, 1644–1655.
Knight, J., Song, L., & Gunatilaka, R. 2009. Subjective well-being and its determinants in rural China. ***China Economic Review***, 20, 635–649.
Konrad, K.A., & Schlesinger, H. 1997. Risk aversion in rent-seeking and rent-augmenting games. ***Economic Journal***, 107, 1671–1683.
Kosfeld, M., & Neckermann, S. 2011. Getting more work for nothing? Symbolic awards and employee performance. ***American Economic Journal: Microeconomics***, 3, 86–99.
Köszegi, B. 2006. Ego utility, overconfidence, and task choice. ***Journal of the European Economic Association,*** 4, 673–707.
Kräkel, M. 2008. Emotions in Tournaments. ***Journal of Economic Behavior and Organization***, *67*, 204-214.
Kuhnen, C. & Tymula, A. 2012. Feedback, Self-Esteem and Performance in Organizations. ***Management Science***, 58, 94–113.
Millner, E.L., & Pratt, M.D. 1991. Risk aversion and rent-seeking: An extension and some experimental evidence, ***Public Choice***, 69, 81–92.
Neckermann, S., Cueni, R., & Frey, B. 2014. Awards at work. ***Labour Economics***, 31, 205–217.
Niederle, M., & Vesterlund, L. 2007. Do women shy away from competition? Do men compete too much? ***Quarterly Journal of Economics***, 122, 1067–1101.
Sheremeta, R.M. 2011. Contest design: An experimental investigation. ***Economic Inquiry***, 49, 573–590.
Sheremeta, R.M., Masters, W.A., & Cason, T.N. 2012. Winner-take-all and proportional-prize contests: Theory and experimental results. Unpublished manuscript.
Skaperdas, S., & Gan, L. 1995. Risk-aversion in contests. ***Economic Journal***, 105, 951–962.
So, T., Brown, P., Chaudhuri, A., Ryvkin, D., & Cameron, L. 2017. Piece-rates and tournaments: Implications for learning in a cognitively challenging task. ***Journal of Economic Behavior and Organization,*** 142, 11–23.
Tafkov, I. 2013. Private and public relative performance information under different incentive contracts. ***Accounting Review***, 88, 327–350.
Treich, N. 2010. Risk-aversion and prudence in rent-seeking games. ***Public Choice***, 145, 339–349.
Weinberg, B.A. 2009. A Model of overconfidence. ***Pacific Economic Review***, 14, 502–51.

Table 1: Payoff table in a 20-person session

Ranking	Earnings
1	¥11.70
2	¥11.10
3	¥10.50
4	¥9.90
5	¥9.30
6	¥8.70
7	¥8.10
8	¥7.50
9	¥6.90
10	¥6.30
11	¥5.70
12	¥5.10
13	¥4.50
14	¥3.90
15	¥3.30
16	¥2.70
17	¥2.10
18	¥1.50
19	¥0.90
20	¥0.30

Table 2: Summary of Sessions Conducted at Different Locations

Treatment Research Site	C	Private/ fixed pay	Public/ fixed pay	Private/ rank pay	Public/ rank pay
Company X	1	1	1	2	2
Company Y	1	1	1	1	1
Company Z	1	1	1	1	1
University	2	2	2	1	1

Table 3: Participants' demographic characteristics

Variable	**Students (n=220)**		**Employees (n=340)**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Age	19.66	0.76	28.30	10.05
Female	0.56	0.50	0.61	0.49
Married	0.00	0.00	0.36	0.48
Monthly Income	N/A	N/A	1697.57	801.05
Education Level*	5.00	0.00	2.66	0.94

Note: Education Level coding: elementary school =1, junior high=2, senior high=3, college=4, university=5, post-graduate=6.

Table 4: Performance data summary

Variable		**Students (n=220)**		**Employees (n=340)**	
		Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Average per-round productivity	C	10.31	4.44	7.05	2.89
	Private/fixed pay	12.66	3.62	6.65	3.19
	Public/fixed pay	12.02	4.04	7.02	3.26
	Private/rank pay	12.81	3.67	7.08	2.70
	Public/rank pay	11.96	3.41	7.33	3.37
Performance Improvement (*Per_Imp*): Within-person Difference between middle four and first two rounds	C	0.61	1.42	0.63	1.25
	Private/fixed pay	1.50	1.86	0.84	1.17
	Public/fixed pay	1.31	1.46	0.67	1.18
	Private/rank pay	0.99	1.40	0.64	1.35
	Public/rank pay	1.98	1.93	0.71	1.34
Further Improvement (*Fur_Imp*): Within-person Difference between last two and the middle four rounds	C	–0.18	1.55	–0.30	1.31
	Private/fixed pay	0.66	2.19	0.43	1.14
	Public/fixed pay	0.35	1.35	0.59	1.06
	Private/rank pay	0.44	2.10	0.23	1.43
	Public/rank pay	–0.18	3.13	0.19	1.49
Risk Aversion		2.88	3.07	3.87	3.25

Table 5: Regression results and Chi² tests for Performance Improvement (*Per_Imp*)

Panel A: Regression Coefficients

	Student Data (n=220)		**Employee Data (n=340)**	
	Complete Model	Model Without RA	Complete Model	Model Without RA
T2: private/fixed pay	1.79***	0.90***	0.76**	0.21
	(0.42)	(0.33)	(0.33)	(0.23)
T3: public/fixed pay	1.70***	0.70**	0.53*	0.04
	(0.44)	(0.33)	(0.33)	(0.23)
T4: private/rank pay	1.08**	0.39	0.62**	0.01
	(0.53)	(0.39)	(0.30)	(0.21)
T5: public/rank pay	2.49***	1.37***	0.40	0.08
	(0.51)	(0.39)	(0.29)	(0.21)
Risk Aversion	0.11		−0.06	
(*RA*)	(0.07)		(0.04)	
RA×T2	−0.34***		−0.13**	
	(0.10)		(0.06)	
RA×T3	−0.31***		−0.12**	
	(0.10)		(0.06)	
RA×T4	−0.20*		−0.16***	
	(0.11)		(0.06)	
RA×T5	−0.39***		−0.12**	
	(0.12)		(0.06)	
Constant	0.23	0.60**	0.92***	0.63***
	(0.34)	(0.25)	(0.23)	(0.16)
Panel B: Chi² tests (1 df for tests 1-4 and 3 df for test 5)				
	Complete Model	Model Without RA	Complete Model	Model Without RA
1. *T3−T2* =0.	0.06, *p* =0.81	0.44, *p* =0.51	0.48, *p* =0.49	0.56, *p* =0.46
2. *T4−T2* =0.	2.26, *p* =0.13	2.02, *p* =0.16	0.21, *p* =0.64	0.82, *p*=0.37
3. *T5−T3*=0.	2.94, *p*=0.09	3.42, *p*=0.06	0.22, *p*=0.64	0.05, p =0.83
4. *T5−T4* =0.	6.69, *p*=0.01	5.58, *p*=0.02	0.79, *p*=0.37	0.12, *p*=0.73
5. *RA×T2=RA×T3=RA×T4=RA×T5*	2.78, *p*=0.43		0.62, *p*=0.89	

Notes: The dependent variable in Panel A is performance improvement (*Per_Imp)* from the first two rounds to the middle four rounds. A maximum-likelihood random-effect specification is used to control for unobserved session effects. Standard errors are in parentheses. All *p*-values are two-tailed. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 6: Regression results and Chi^2 tests for Further Improvement (*Fur_Imp*)

Panel A: Regression Coefficients

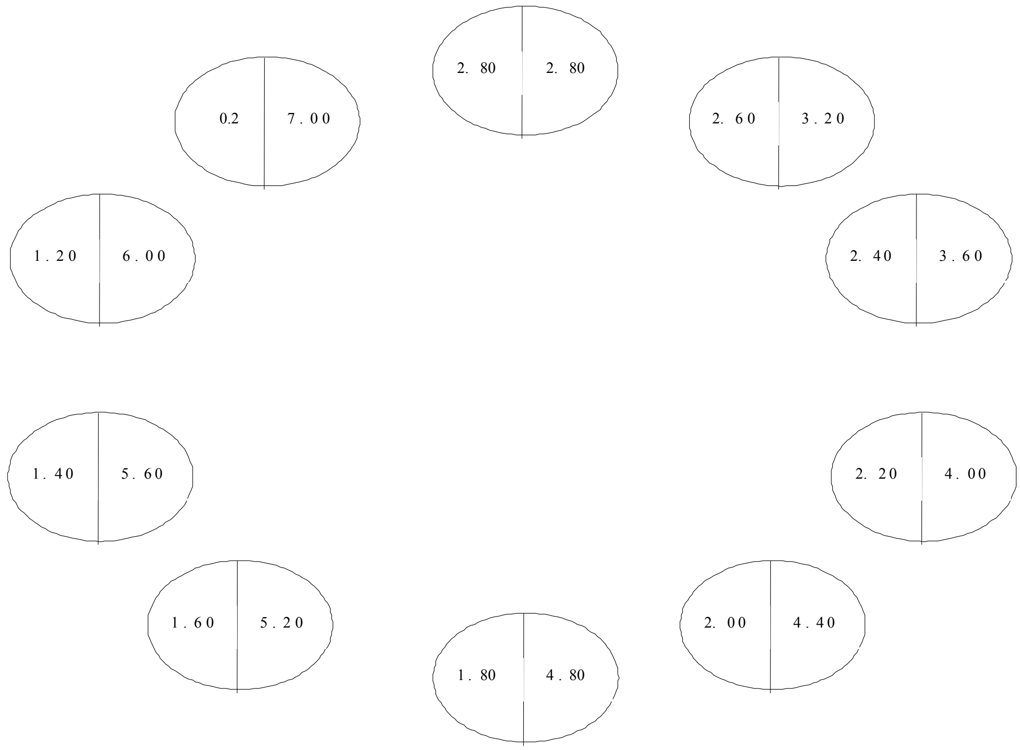
	Student Data (n=220)		**Employee Data (n=340)**	
	Model 1	Model 2	Model 1	Model 2
PER_IMP	−0.15*	−0.17**	−0.25***	−0.23***
	(0.09)	(0.08)	(0.06)	(0.06)
FEED (*T2*, *T3*, *T4* or *T5*)		0.72**		0.65***
		P=0.047		(0.18)
		(0.36)		
T2: private/fixed pay	0.86		0.50	
	(0.58)		(0.38)	
T3: public/fixed pay	0.92		0.63	
	(0.59)		(0.38)	
T4: private/rank pay	0.70		0.44	
	(0.70)		(0.35)	
T5: public/rank pay	0.98		0.33	
	(0.70)		(0.34)	
Risk Aversion	0.08		−0.05	
(*RA*)	(0.10)		(0.05)	
RA×T2	0.10		0.07	
	(0.14)		(0.07)	
RA×T3	−0.08		0.06	
	(0.13)		(0.07)	
RA×T4	−0.01		0.02	
	(0.14)		(0.07)	
RA×T5	−0.26		0.04	
	(0.16)		(0.07)	
Constant	−0.36	−0.08	0.08	0.15
	(0.45)	(0.32)	(0.27)	(0.17)

Panel B: Chi^2 tests (3 df)

	Student Data	Employee Data
T2=T3=T4=T5	0.17, p =0.98	0.78, p =0.85

Notes: The dependent variable in Panel A is further improvement (*Fur*_Imp) from the middle four rounds to the last two rounds. A maximum-likelihood random-effect specification is used to control for unobserved session effects. Standard errors are in parentheses. All p-values are two-tailed. ***, ** and * denote significance at 1%, 5% and 10% respectively.

FIGURE 1
Risk Attitude Measure



Experimental Instructions

Your Participant Number is __________.

Thank you for participating today. All of your responses in this study will remain completely anonymous. It is important that during this experiment you do not talk or make any noise that might disrupt others around you. If you have any questions, please raise your hand and the experimenter will answer your questions individually.

During this experiment you will be asked to add up sets of five double-digit integers. There will be nine rounds in which you will be given a number of such sets of five integers. The first round is a trial round for you to get familiar with the task while the remaining eight rounds are experimental rounds, which will be used to calculate your earnings. You are not allowed to use a calculator, but you may write numbers down on scratch paper provided by us. The numbers are randomly drawn and each problem is presented in the following way.

98	42	69	50	78	

You will have a *Workbook* that will contain all of your work. Your job is to solve as many problems as you can in each round. Each round lasts 3 minutes. Your earnings in this experiment will depend on your performance and/or the specific compensation method applied to each of the eight experimental rounds. Once we begin the experiment, you will not be able to look ahead to future pages or to go back to previous pages.

To ensure anonymity, please write down only your participant number on each page of the *Workbook*. Please do not write your name on any of these materials. The data will only be identified by the participant code assigned to you and will not at any point be connected to your name or face in any way.

Please make sure that you completely understand the instructions for the experiment. Once again, remember not to make any noises that might disturb others around you. If you have any questions, raise your hand and we will answer your questions individually.

Control: You will earn RMB 6.00 in this round, regardless of the number of words you create in this round.

Private/Fixed Pay: You will earn RMB 6.00 in this round, regardless of the number of words you create in this round. In addition, your performance will be ranked against everyone else in the session and your ranking information will be reported to you privately.

Public/Fixed Pay: You will earn RMB 6.00 in this round, regardless of the number of words you create in this round. In addition, your performance will be ranked against everyone else in the session and everyone's ranking information will be linked to his/her name and will be reported to everyone publicly.

Private/Rank Pay: Your performance will be ranked against everyone else in the session and your ranking information will be reported to you privately. You will be paid based on your performance ranking,

Public/Rank Pay: Your performance will be ranked against everyone else in the session and everyone's ranking information will be linked to his/her name and will be reported to everyone publicly. You will be paid based on your performance ranking,

Sample Workbook Page

						Answer							Answer
Line 1	69	95	12	72	25		**Line 21**	33	55	40	65	48	
Line 2	95	36	77	85	50		**Line 22**	37	79	88	21	64	
Line 3	80	82	55	24	31		**Line 23**	12	38	12	48	49	
Line 4	65	72	97	87	74		**Line 24**	41	79	33	96	60	
Line 5	25	30	12	72	97		**Line 25**	18	44	68	11	34	
Line 6	83	49	47	37	49		**Line 26**	38	54	83	64	97	
Line 7	30	93	74	71	44		**Line 27**	81	27	85	31	87	
Line 8	87	80	14	17	27		**Line 28**	37	77	21	92	84	
Line 9	51	27	71	76	63		**Line 29**	43	87	83	32	59	
Line 10	31	41	40	10	19		**Line 30**	48	73	94	75	35	
Line 11	15	17	76	46	30		**Line 31**	44	65	79	81	69	
Line 12	68	87	98	49	37		**Line 32**	63	98	72	46	64	
Line 13	14	74	50	85	50		**Line 33**	94	73	54	12	13	
Line 14	25	15	15	10	92		**Line 34**	81	36	43	88	71	
Line 15	20	13	88	22	37		**Line 35**	83	99	38	20	35	
Line 16	59	42	99	50	81		**Line 36**	19	11	99	44	53	
Line 17	48	31	33	15	14		**Line 37**	80	74	91	55	77	
Line 18	56	14	10	77	17		**Line 38**	84	60	55	49	10	
Line 19	60	96	44	33	91		**Line 39**	16	90	41	82	25	
Line 20	86	83	65	47	67		**Line 40**	45	66	14	84	41	