TURNING UP THE VOLUME:
AN EXPERIMENTAL INVESTIGATION OF THE ROLE OF MUTUAL MONITORING IN TOURNAMENTS

R. Lynn Hannan
J. Mack Robinson College of Business
Georgia State University
P.O. Box 3989
Atlanta, GA 30302
rhannan@gsu.edu

Kristy L. Towry
Goizueta Business School
Emory University
1300 Clifton Rd.
Atlanta, GA 30322
kristy_towry@bus.emory.edu

Yue (May) Zhang
College of Business Administration
Northeastern University
360 Huntington Ave.
Boston, MA 02115
yu.zhang@neu.edu

July, 2011

We thank Jan Bouwens, Jason Brown, Eddy Cardinaels, Joe Fisher, Theresa Libby, Laureen Maines, Bill Messier, Chad Simon, Jason Smith, Arnie Wright, participants at the Oklahoma Behavioral Accounting Conference and the Global Management Accounting Research Symposium, and workshop participants at Florida State University, George Mason University, Indiana University, Katholieke Universiteit Leuven, Tilburg University, and University of Nevada - Las Vegas for providing helpful comments on an earlier version of this paper. An earlier study, which resulted in the current paper, benefited from comments provided by Jon Davis, Ranjani Krishnan, Brian Mayhew, workshop participants at Georgia State University, University of Wisconsin and discussions at the 2009 Management Accounting Section Research Conference, and the 2009 American Accounting Association Annual Meeting.

TURNING UP THE VOLUME: AN EXPERIMENTAL INVESTIGATION OF THE ROLE OF MUTUAL MONITORING IN TOURNAMENTS

ABSTRACT

This study investigates experimentally how mutual monitoring affects effort when individuals are compensated via rank-order tournaments. The prior literature generally assumes that mutual monitoring will have a detrimental effect on effort in tournaments, because it will facilitate collusion among tournament participants. Our two experimental studies show that this is not necessarily the case. In our first study, we find that mutual monitoring actually increases effort, because participants do not attempt to collude but rather behave competitively. In our second study, we manipulate the mindset of participants at two levels, collusive and competitive. We develop theory to predict that mutual monitoring leads to lower (higher) effort when employees enter the tournament with a collusive (competitive) mindset. In other words, mutual monitoring "turns up the volume" of mindset effects. Consistent with our predictions, mutual monitoring leads to lower effort when participants have a collusive mindset and (eventually) higher effort when they have a competitive mindset. Overall, our study suggests that, although collusion is generally viewed as a potential threat to the efficacy of tournament contracts, the threat is only realized under specific circumstances. As such, our study provides an explanation for the widespread use of tournament contracts, despite the potential for collusion discussed extensively in the prior literature. Importantly, our findings also suggest that allowing employees to observe each other's productive effort in tournament incentive settings may have positive or negative effects for the firm, depending on whether environmental factors create a collusive or competitive mindset.

I. INTRODUCTION

Over the past several decades, rank order tournaments (compensation schemes in which individuals are ranked on an ordinal scale based on performance and compensated according to their ranks) have received considerable attention in the agency theory-based literature. The use of tournament incentive schemes is pervasive, with surveys revealing that as many as one-fourth of Fortune 500 companies and one-third of U.S. corporations use such "forced ranking" systems (Boyle 2001; McGregor 2006). Potential benefits of tournaments include a selection effect, via which tournaments attract skilled employees willing to compete, and an incentive effect, via which tournaments motivate high effort and performance (Prendergast 1999; Lazear 2000). The positive effects of tournaments, however, may be reduced or even eliminated if employees collude to minimize effort, a concern that pervades the discussion of tournaments (e.g., Dye 1984, Gibbons and Murphy 1990, Prendergast 1999, Harbring 2006).

Scholars have suggested that collusion will be particularly problematic when employees can mutually monitor one another. (Mutual monitoring is the ability of peers to observe each other's productive activities, and thereby obtain private information not available to their superior.) For example, Holmstrom and Milgrom (1990 p. 86) observe that mutual monitoring facilitates collusion such that multiple agents theoretically behave like a single agent. Gibbons and Murphy (1990) recognize the importance of mutual monitoring when they argue that collusion is likely to be a greater threat in the workplace than in the product market, because the workplace facilitates mutual monitoring whereas the product market does not. Mutual monitoring is believed to facilitate collusion by enhancing enforcement of collusive agreements via the ability to retaliate against reneging. Based on this reasoning, conventional wisdom suggests that mutual monitoring will reduce effort in tournaments.

However, empirical evidence on the effect of mutual monitoring in tournaments is scant, and there are reasons – both theoretical and anecdotal – to question whether the conventional wisdom will prevail. From a theoretical perspective, it is possible that mutual monitoring will facilitate competition rather than collusion, by encouraging social comparison and allowing tournament participants to assess the level of effort necessary to win the tournament. That is, mutual monitoring allows tournament participants to observe competitors' effort levels and to increase their own effort levels in response, thereby ratcheting up effort levels. From an anecdotal perspective, tournaments are often used in environments where mutual monitoring is possible, which is inconsistent with the notion that the potential for collusion is heightened in such settings. For example, retailer AnnTaylor uses a tournament scheme as a motivating tool for its retail store employees, and actually encourages mutual monitoring, by providing headsets through which individual sales associates' performance can be communicated throughout the store (O'Connell 2008). In summary, both theoretical arguments and the common use of tournaments in conjunction with mutual monitoring call into question the intuition that mutual monitoring facilitates collusion in tournaments. In this paper, we test this intuition, examining whether and by what processes mutual monitoring affects effort in tournaments.

We conduct this research using two experiments. In Study 1, participants act as employees who are compensated through a tournament scheme based on performance, which is a function of their chosen effort levels as well as a noise parameter. The independent variable is mutual monitoring (yes or no), which we manipulate by allowing (or not allowing) participants to directly observe each other's effort levels. In this setting, we find no evidence that mutual monitoring facilitates collusion. In fact, we find that mutual monitoring actually increases effort. Further examination of the data reveals that very few of the participants in this experiment even attempted to collude, regardless of whether or not they had the ability to mutually monitor one another. Rather, it appears that the information gathered through mutual monitoring was used by participants to compete rather than to collude.

The results of Study 1 lead to a more refined theory regarding the role of mutual monitoring in tournaments. We conjecture that if participants in Study 1 had been inclined to collude and had attempted to do so, then the ability to mutually monitor one another might have increased the success of their collusive attempts, thereby decreasing effort. Because they were instead inclined to compete, mutual monitoring reinforced their competitive behavior. In other words, our refined theory suggests that mutual monitoring will not have the uni-directional (negative) effect on effort previously presumed. Rather, it will "turn up the volume" on whatever mindset with which participants enter the tournament. If this theory is correct, then the effect of mutual monitoring will depend on the mindset with which participants enter the tournament. This theory motivates our second experiment.

The experimental setting for Study 2 is similar to that of Study 1. However, in this experiment, we use a 2 x 2 between-subjects experimental design. The first independent variable is mutual monitoring, manipulated as in the first experiment. The second independent variable is the mindset (collusive versus competitive) with which participants enter the tournament. We manipulate this variable by emphasizing through the instructions either the collusive or the competitive strategy for participating in the tournament. We find that, as expected, mutual monitoring *decreases* effort when participants enter the tournament with a collusive mindset. In contrast, by the final period of interaction, mutual monitoring *increases* effort when participants enter the tournament with competitive mindsets. Further analysis of the data suggests that, consistent with our conjecture, mutual monitoring reinforced competitiveness, leading to higher effort.

Overall, the results from these two studies suggest that mutual monitoring is unlikely to be detrimental to effort in tournaments unless employees enter the tournament with some inclination to collude. In fact, mutual monitoring can increase effort if employees are inclined to compete. Although collusion is generally viewed as a potential threat to the efficacy of tournament contracts, our results suggest that the threat is only realized under specific circumstances. As such, our study provides an explanation for the widespread use of tournament contracts despite the potential for collusion discussed extensively in the prior literature. Importantly, our findings also suggest that allowing employees to observe each other's productive effort in tournament incentive settings may have positive or negative effects for the firm, depending on whether environmental factors create a collusive or competitive mindset.

This study contributes to the broad stream of accounting research that investigates empirically the effectiveness of formal control mechanisms in general, and tournaments and mutual monitoring in particular. Prior research has considered the role of mutual monitoring on optimal contracting in a variety of other incentive contexts (e.g., Arya, Fellingham and Glover 1997, Ma 1988, Towry 2003), and has generally viewed mutual monitoring as a control tool, which can be used as an additional lever by which to induce employee effort. In this paper, we integrate the literature on mutual monitoring with that on tournament contracts to investigate how the ability of employees to observe their peers affects effort in a tournament setting.

In demonstrating that the effect of mutual monitoring in tournaments depends on the mindset with which employees enter the tournament, our research speaks to the broader implication that various management control devices must be considered as parts of an overall system of control. While prior research has considered the use of both tournaments and mutual monitoring in optimal contracting solutions, our research suggests that the effects of such formal control devices likely interact in complex ways.

The remainder of this paper is organized as follows: the next section reviews the literature, develops the initial hypothesis, and describes Study 1. Section III provides additional theory, develops the refined hypotheses, and describes Study 2. Section IV discusses the results and concludes this paper.

II: STUDY 1

Tournaments

This study investigates how mutual monitoring affects effort in rank-order tournaments. A rank-order tournament is a specific type of relative performance incentive system, in which individuals are performance-ranked on an ordinal scale and compensated according to their ranks, with a fixed compensation for each rank. The compensation is not a function of either the performance differential or actual performance (Prendergast 1999; Bull, Schotter and Weigelt 1987).[1] Such schemes are frequently observed in the corporate world (McGregor 2006; Lambert, Larcker and Weigelt 1993). Consider, for example, periodic sales contests, in which the highest performing personnel receive a bonus or a prize, as well as the implicit tournaments created by the competition for promotions (Green and Stokey 1983, Orrison, Schotter and Weigelt 2004). Tournaments also take the form of the notorious "rank and yank" systems, in which the lowest ranked employees lose their jobs (Grote 2005). This type of tournament has been advocated by well-known companies such as General Electric, American Express, Hewlett Packard, and Goldman Sachs (Boyle 2001).

[1] Although other types of tournament schemes exist in practice, such as linear relative performance evaluation schemes or tournaments combined with piece rate schemes, we restrict our study to rank order tournaments in which compensation is determined entirely by the ordinal rank.

Tournament incentive schemes can be effective at motivating effort because employees compete to be ranked higher than others. A key advantage of tournaments is that they can motivate effort at a lower cost than individual performance contracts because tournaments help filter out the environmental uncertainty that is common to all employees. This reduced environmental uncertainty results in lower risk being imposed on the employees, and thus lower average required compensation (Lazear and Rosen 1981). In an optimally designed tournament, economic theory predicts that employees will behave competitively with respect to effort, and the Nash equilibrium is for all employees to provide effort at the level where the marginal cost of effort is equal to the marginal expected return. At this equilibrium, all employees will choose the same effort level, and so the outcome of the tournament is essentially a lottery, in which each employee has an equal probability of winning.[2]

However, employees can improve over this equilibrium outcome if they collude (Dye 1984; Potters *et al*. 2004). Specifically, the Pareto optimal solution from the employees' perspective is for all employees to provide minimal effort. If all employees do so, the outcome of the tournament is once again a lottery, in which each employee has an equal probability of winning. Therefore, the expected probability of winning the tournament is the same under both solutions; however, the disutility of effort is lower under the Pareto optimal solution. While this collusive outcome is Pareto optimal to the employees, collusive behavior is not an equilibrium. That is, as long as other employees provide minimal effort, any individual employee can obtain a dramatic increase in ranking and incentive pay by increasing effort only slightly (up to the point where the marginal cost of effort equals the marginal expected return).

[2] This analysis assumes that skill levels and utility functions (e.g., risk preferences, disutility for effort) are equal across employees.

Basic Setting

In this section, we describe the basic setting used to develop and test our theory, which is related to several analytic models that have examined rank-order tournament incentive schemes (Lazear and Rosen 1981, Nalebuff and Stiglitz 1983, Prendergast 1999; Orrison *et al*. 2004). The simple model consists of two identical agents i ($i = 1, 2$), competing for two prizes, M or m, with M > m > 0 (Orrison *et al*. 2004). For simplicity, let agent i's output y_i equal $e_i + \varepsilon_i$, where e_i represents agent i's effort, and ε_i is a random factor, distributed uniformly over the interval [-a, a] (a > 0) and independently across agents. Following Bull *et al*. (1987), the agents' utility functions are assumed to be separable into the payment (p) received and the effort (e) exerted (i.e., $U_i = w(p_i) - c(e_i)$), where $w(p_i)$ is the utility from the payment received, and $c(e_i)$ is the disutility of effort. In our experimental setting, we adopt Bull *et al*.'s (1987) operationalization of the two components of utility, setting $w(p_i) = p_i$ and $c(e_i) = e_i^2 / c$. Denote $\rho(e_i, e_{-i})$ as agent i's probability of winning M, given any vector $\mathbf{e} = (e_1, e_2)$ of effort choices by the agents. Agent i's expected utility from such a choice is $EU_i\ (e_i, e_{-i}) = m + \rho(e_i, e_{-i})[M - m] - c(e_i)$. At the unique pure strategy Nash equilibrium, the optimal effort level, $e^* = [(M - m) * c] / 4a$ (Bull *et al*. 1987).

In our studies, participants earn points, which are converted to dollars for their cash payment. The winner's final prize, M, equals 200 points and the loser's final prize, m, equals 100 points. In order to compare our results to the analytic models, we operationalize effort via a chosen effort level and cost of effort via point deductions. Specifically, the effort level, e, ranges from 1 to 40 and the cost of effort is determined as $e_i^2 / 7$. The random shock is uniformly distributed over the interval [-10, 10] (i.e., c = 7, e = a whole number from 1 to 40, and a = 10). Based on these parameters, the Nash equilibrium effort level equals 17.5 (i.e., e* = [(200-100) *

7] / (4*10) = 17.5). On the other hand, the Pareto optimal outcome is for the agents to collude, each selecting the minimum effort level of one.

Mutual Monitoring

The primary purpose of this study is to investigate how mutual monitoring affects effort under tournaments. Mutual monitoring is the ability of peers to observe each other's productive activities, resulting in asymmetric information between the peers and their superior. The degree of mutual monitoring varies in the workplace due to such factors as differential levels of interaction, adjacency in work locations, or access to data from information systems.[3] For example, employees working on a production line in a traditional manufacturing setting may have near perfect information on the actions of their co-workers. On the other hand, in the increasingly common "virtual workplace," employees will likely know even less than the central office about their co-workers' activities. This wide variety suggests that it is important to understand the effect of mutual monitoring and its interaction with various other control mechanisms.

The ability of employees to mutually monitor one another is often interpreted as providing an additional lever by which managers can improve the control environment in non-

[3] A significant body of literature in accounting (e.g., Frederickson 1992; Hannan, Krishnan and Newman 2008; and Tafkov 2010) addresses the effect of providing relative performance information to employees. There are two important distinctions between the relative performance information and the mutual monitoring literatures. First, the mutual monitoring literature addresses the ability of employees to observe the actions and/or effort of other employees, whereas the relative performance information literature addresses the effect of learning how well one's own performance compares to the performance of others. This distinction implies that the information obtained via mutual monitoring is more informative of actions and/or effort because performance can be affected by both effort and other factors, such as skill or environmental shocks. Second, the mutual monitoring literature addresses the role of monitoring that can be conducted by peers *but not by* management. In fact, the agency theory-based papers on mutual monitoring are predicated on the notion that agents hold information about each other that is not available to the principal, and these papers derive optimal contracts for exploiting this information. In contrast, the investigations of relative performance information typically do not assume information asymmetry between the principal and agents. Consistent with these distinctions, the participants who receive relative performance information in Hannan *et al*. (2008) only learn their own performance ranking, from which the effort and actions of others can only be inferred with noise, and the relative performance information is used contractually to determine their pay, implying information symmetry.

tournament settings. For example, prior analytic research, based primarily on agency theory, suggests that mutual monitoring can be exploited in the contract design to increase the principal's profit. As described by Towry (2003), there are two general methods by which mutual monitoring can be exploited in contract designs. The first is a vertical system, in which each agent observes other agents' actions and then truthfully reports them to the principal (e.g., Ma 1988). The second is a horizontal system, which relies on team self-management and peer-based control (e.g., peer pressure and the use of tit-for-tat strategies) (e.g., Arya *et al*. 1997).

Empirical research provides evidence that the ability of workers to mutually monitor one another can have positive incentive effects in non-tournament settings. Mas and Moretti (2009), for example, find that low-productivity grocery store checkers increase their effort levels in the presence of high-productivity employees, especially when the employees' production sites and schedules regularly overlap. They interpret this evidence as suggesting that, even in a setting with limited economic incentives, the social pressures induced through mutual monitoring can have powerful incentive effects. Carpenter, Bowles and Gintis (2006) also demonstrate the positive incentive effects of mutual monitoring, providing experimental evidence that mutual monitoring can reduce shirking in public goods games.

Within tournament settings, however, mutual monitoring has been assumed to facilitate collusion (Gibbons and Murphy 1990), resulting in negative consequences for the firm. While theory and evidence regarding mutual monitoring in tournament settings is limited, the industrial organization literature provides relevant analyses when considering the factors that influence the degree of tacit collusion in price setting among competing firms (e.g., Stigler 1964; Green and Porter 1984; Tirole 1988; Kühn 2001; Ivaldi et al. 2003; Motta 2004). Briefly, this literature concludes that such collusion will be more easily sustained when competing firms can observe one another's prices and sales than when they cannot. The intuition is that in order for firms to enforce collusive agreements, it must be possible for them to retaliate against firms who renege, and retaliation is only possible when the firms can mutually observe one another's actions. As described by Ivaldi et al. (2003, p. 22), "the lack of transparency on prices and sales does not necessarily prevent collusion completely, but makes it both more difficult to sustain and more limited in scope."

This basic reasoning on observability across firms can easily be applied to the setting of workplace tournament contracts. In fact, Gibbons and Murphy (1990) argue that collusion will be particularly problematic in the workplace, precisely because the workplace facilitates mutual monitoring. Just as firms need to be able to retaliate against those who renege, employees must also be able to do so in order to enforce collusion. In the absence of mutual monitoring (or with weak monitoring), it would be tempting for another employee to renege on a collusive agreement by putting in high effort and attributing the resultant high performance to luck because it would be difficult to verify the veracity of such a claim. As a result, attempts to collude are likely to unravel quickly. With mutual monitoring, on the other hand, employees can employ a tit-for-tat strategy by continuing to collude as long as others do so, and to defect in response to another employee's defection.[4] As a result, mutual monitoring facilitates collusion by providing information useful for enforcing collusive agreements.

To our knowledge, only one study has investigated whether mutual monitoring will indeed result in collusion, albeit in a relative compensation rather than a tournament setting. Bandiera, Barankay and Rasul (2005) analyze personnel data from a British fruit farm and find

---

[4] Although the Nash equilibrium in *finitely*-repeated games is for non-cooperation in all stages, experiments show that cooperation emerges even when interaction is not infinite (e.g., Selten and Stoecker 1986). Kreps *et al*. (1982) show analytically that infinite interaction is not necessary as long as players believe others will use a tit-for-tat strategy.

that when workers are able to observe others' activities, productivity is significantly lower under a relative incentive system compared to a piece-rate compensation system. The authors argue that under the relative compensation system, each worker's effort imposes a negative externality on their co-workers' pay. Therefore, workers have incentives to form an implicit collusive agreement to reduce common productivity. However, it appears to be difficult to sustain such collusive agreement when the workers are not able to perfectly monitor each other by working alongside their peers. Based on this reasoning, one might expect mutual monitoring to decrease effort in tournaments.

On the other hand, employees may use the information obtained via mutual monitoring to compete rather than to collude. If this is the case, then the effect of mutual monitoring in tournaments is uncertain in that it may increase or have no effect on effort. First, assuming that participants have utility for only wealth and leisure, and that they behave as predicted by economic theory, tournament participants would choose the Nash equilibrium effort level regardless of mutual monitoring, suggesting that mutual monitoring will have no effect on efforts in tournament. Second, from a behavioral perspective, mutual monitoring could lead to higher effort in tournaments by increasing the salience of social comparison. Specifically, social comparison theory (Festinger 1954) posits that individuals have a drive to continually compare themselves with others in order to evaluate their own abilities. When individuals perform worse than others, self identity suffers (Tesser and Campbell 1980). Thus, mutual monitoring might encourage individuals to increase effort in order to win the tournament and maintain a positive self image. Moreover, assuming cognitive limitations, mutual monitoring allows tournament participants to use other participants' effort levels as benchmarks for choosing their own effort levels, ratcheting effort up to the point where the marginal cost of effort equals the marginal benefit, or even higher (if participants have utility for winning).

Based on the above discussion, it is unclear whether mutual monitoring will decrease effort in tournaments, as has been assumed in the prior literature. Because theory does not allow for an unambiguous directional prediction, we state the first hypothesis in the null form.

H1: Effort will not differ whether or not employees can mutually monitor one another.

Experimental Design

We conduct an experiment in which dyads work under a tournament contract for points, which are converted to dollars at the end of the session.[5] The primary experimental design is a one-way between-subjects design, in which we manipulate mutual monitoring. Dyads remain paired for four periods; thus, period is a third, within-subjects, factor. Mutual monitoring is manipulated by allowing each participant to observe (or not observe) the other employee's effort level. The primary dependent variable is effort. We use an abstract measure of effort, by having the participants choose an effort level between 1 and 40. Effort is costly (i.e., points are deducted for the level of effort chosen), and expected performance increases in effort.[6] Specifically, the cost of effort is a convex function computed as $\text{effort}^2 / 7$. A key advantage of operationalizing the effort choice and the cost of effort in this abstract manner is that it allows us to calculate and compare decisions to economic benchmarks based on game theoretic reasoning (i.e., the Nash equilibrium and the Pareto optimal outcome).

---

[5] While in the real world, tournaments likely involve more than two employees, the theory we test does not rely on more than two employees. Thus, we are able to efficiently test the theory by limiting participation to two employees per tournament.

[6] As discussed by Baiman (1982), agency theory defines effort as an action that is controlled by the agent, results in disutility, and is correlated positively with output. Our abstract effort choice satisfies these criteria. Experimental studies in accounting that have employed this abstract method include Frederickson (1992), Frederickson and Waller (2005), Hannan (2005) and Kuang and Moser (2009).

Each employee's performance is calculated as effort plus a random shock. The shock is determined randomly from the uniformly-distributed set of {-10, 10} and is independent across participants and periods. The purpose of the random shock is to limit the ability of one employee to infer the other employee's actual effort level, given the observed performance ranks. Thus, this feature is essential for us to operationalize our "no mutual monitoring" condition.[7]

Compensation in each period is based on the participants' performance ranks in the dyad. The participant with the higher ranked performance receives pay of 200 points and the participant with the lower ranked performance receives pay of 100 points. In the case of tied performance, the computer randomly determines the winner for compensation purposes.

Participants in all conditions learn their own performance levels, ranks, and points at the end of each period. In the mutual monitoring condition, each participant also learns the actual effort level of the other employee. Therefore, mutual monitoring provides perfect information about actions of the other employee, and participants can use this information to plan responses (e.g., increase effort to become more competitive or punish the other employee for defecting from a collusive agreement).[8] In contrast, in the no mutual monitoring condition, information about effort choices of the other employee is quite noisy, because each participant only learns the other employee's performance rank, and this ranking is affected by a substantial random shock (as much as +/- 10 on an input base of 1 to 40).

---

[7] Although an advantage of tournaments relative to individual performance contracts is their ability to filter out common environmental uncertainty, such filtering is not central to our theory, which relates to comparisons *within a* tournament rather than *across* tournaments and other types of compensation contracts. To simplify the experiment, we hold common environmental uncertainty at zero in our experimental tournament.

[8] The effort level information provided to the employees is an abstraction of the ability to mutually monitor. That is, just as the experiment uses an abstraction of effort by having participants choose a costly effort level, it also uses an abstraction of mutual monitoring by informing the other participant of the actual effort level chosen. Importantly, this information is conceptually derived from mutual monitoring and not from an information system that the firm could exploit contractually to force an effort level.

We adapted the parameters for our experiment from Bull *et al*. (1987). As described earlier, the unique pure strategy Nash equilibrium is for each participant to provide an effort level of 17.5, whereas the Pareto optimal outcome is for each to choose the minimal allowed effort level of one.

Participants and Procedures

Participants are undergraduate business students from a large U.S. university.[9] Two experimental sessions were conducted, one for each experimental cell. The mutual monitoring condition had eight participants, and the no mutual monitoring condition had ten participants. Participants interact over a computer network,[10] with anonymity preserved both during and after the experimental sessions.

As shown in Figure 1, each experimental session has four cycles, consisting of four periods each. Participants are matched as dyads and interact for four periods before being matched into new dyads for a new cycle of four periods. Our matching protocol uses a "turnpike design," such that no participant is paired with the same person twice or with any person who has been paired with anyone the participant has interacted with previously. This design ensures that actions taken by a dyad cannot influence the behavior of any participant with whom they will be matched in the future (Cooper *et al.* 1996). Thus, each unique dyad represents an independent observation.[11] This procedure results in 20 independent observations (four cycles x five dyads per cycle) for the session with ten participants and 16 independent observations (four cycles x four dyads per cycle) for the session with eight participants. Further, we capture four periods of repeated interaction for each unique dyad.

[9] The average age of participants is 20 years. They report 1.1 years of work experience, on average. Fifty-six percent of the participants are male.

[10] The experiment was programmed in Z-Tree (Fischbacher, 2007).

[11] As a robustness check, we also test our hypotheses within each cycle, and find inferentially identical results.

At the beginning of the experiment, participants are provided with instructions and these are read aloud by one of the authors. Participants are told that, to motivate them to work hard, the company they work for uses an incentive compensation scheme based on how well they perform relative to a partner. The tournament incentive scheme is described, including how effort relates to performance and the random noise parameters. Participants are provided with a table relating effort levels to their respective costs, a copy of which is reproduced as Table 1. Participants are informed of the matching protocol, i.e., that they will interact with the same partner for four periods, after which they will be matched with a new partner who has never been matched with a previous partner. They are also informed that the session will last for a total of four cycles and that one period will be selected at random for payment. After reading the instructions, participants take a quiz to ensure that they understand the instructions. All participants are required to score 100 percent on the quiz before the first period commences.

The procedures in each period are the same for both conditions. Each period begins with a chat box appearing on each participant's computer screen, which participants may use to communicate with their dyad partner. Communication is unrestricted except that participants are not permitted to provide any information that would reveal their personal identities. We allow this communication in order to provide a means for the participants to discuss collusion if they so desire. After 45 seconds, the chat session ends, a new screen appears, and each participant privately inputs the effort level for the period. After all participants have inputted their effort levels, a new screen provides participants with the results for the period. Participants in all conditions are shown their own effort levels, cost of effort, performance, tournament ranking, and net points for the period. Participants in the mutual monitoring condition are also shown the effort level chosen by the other employee in their dyad.

The same procedures are followed for a total of four periods to complete a cycle. At the beginning of a new cycle, participants are matched with a new partner via the turnpike design described above, and this new dyad interacts for four periods. This process continues for a total of four cycles so that at the conclusion of the session, each participant has completed four four-period cycles, for a total of sixteen periods.

At the end of the experiment, participants complete a post-experimental questionnaire. One period is randomly selected to be the payment period. Participants receive a cash payment for this period at the rate of one dollar for every ten points in addition to a participation fee of $5.00.

Results

H1 considers whether mutual monitoring will affect effort in tournaments. As shown in Table 2, the results contradict the conventional wisdom, because the evidence suggests that mutual monitoring does not decrease effort in our setting. In fact, mean effort is higher with mutual monitoring (22.95) than without mutual monitoring (18.73) ($F = 6.90$, $p < 0.01$, two-tailed).[12] An examination of the post-experimental questionnaire responses provides insight into this result. Specifically, one question asked participants, "Did you try to agree on an effort level with the other employee?" Only three of the eighteen participants responded yes to this question, indicating that most participants were not inclined toward collusion, and the ability to mutually monitor one another did not increase that inclination. We can conclude from these results that mutual monitoring does not increase the *incidence* of collusive attempts. However, given the low level of collusive attempts observed, whether mutual monitoring increases the *success* of collusive attempts is still an open question.

[12] Recall that each dyad represents an independent observation under the turnpike design used for our matching protocol. Thus, our dependent measure is the average effort for the two participants in each unique dyad across the four periods in which the dyad interacted.

It is notable that mutual monitoring actually increased effort (rather than simply failing to decrease it) in Study 1. In fact, with mutual monitoring, the average effort level is significantly higher than the Nash equilibrium of 17.5 ($t = 3.76$, $p < 0.01$ two-tailed), whereas without mutual monitoring, effort is indistinguishable from the Nash equilibrium ($t = 1.45$, $p = 0.16$ two-tailed).[13] We conjecture that the positive effect of mutual monitoring on effort is the result of participants being predisposed towards competition, and mutual monitoring reinforcing that competition. Responses to a post-experimental question provide evidence that participants viewed the tournament competitively. Specifically, the mean response to "how badly to you want to win the bonus" is 6.1 on a 7.0 scale. However, the process by which mutual monitoring leads to higher effort is not clear. We do not observe any trends in effort across time, and so we have no evidence of the ratcheting up that might occur if participants are using other participants' effort levels as benchmarks, and then responding by increasing their own effort levels. The result is consistent with social comparison theory (Festinger 1954) and with prior experimental (e.g., Falk and Ichino 2006, Hannan et al. 2008, Tafkov 2010) and field research (e.g., Mas and Moretti 2009), which has shown that less productive employees work harder in the presence of more productive employees, even when there is no economic incentive to do so. We conjecture that mutual monitoring makes the presence of others more salient, which increases the effects of social comparisons and increases effort. In the following section, we consider this possibility, as we use the results of Study 1 to refine our theory, and describe Study 2, which empirically tests the refined theory.

III. STUDY 2

Employee Mindset

[13] This is consistent with prior research, which indicates that individuals experience positive utility from winning, in addition to any utility associated with pecuniary awards (Ku, Malhotra and Murnighan 2005).

In building the refined theory, we begin by introducing the psychological construct of mindset. The term mindset has been used in many contexts, and represents the categorization of cognitive states or processes by such descriptors as deliberative versus implemental, abstract versus concrete, local versus global, or approach versus avoidance (Gollwitzer and Bayer 1999; Förster, Liberman and Friedman 2009). Our use of the term is more general, in that we define mindset as the general inclination or strategic focus with which an employee enters the tournament: collusive versus competitive. This definition is consistent with Harinck and De Dreu's (2008) use of the term in a negotiation context, in which they describe competitive versus cooperative mindsets as reflecting the strategic focus with which individuals enter into a negotiation.[14] In the tournament context, however, we replace the term "cooperative" mindset with "collusive" mindset, because cooperation that is detrimental to the principal is typically called collusion (Holmstrom and Milgrom 1990).

An employee's mindset can be influenced by dispositional (Bogaert, Boone and Declerck 2008) and situational factors (De Dreu *et al*. 2006, p. 928) as well as cultural values (Hofstede 1980, 2001). Importantly, an employee's mindset may also be influenced by managerial decisions that collectively determine the culture of the firm. For example, both experimental (Zhang 2008) and field (Chen and Sandino 2010) studies suggest that management practices that are perceived as unfair can lead employees to develop a collusive mindset. Relatedly, many firms enact programs and policies that are explicitly designed to build social ties among employees. These social ties can have a powerful effect on decision-making. Research on social

[14] A related but more specific construct is goal orientation (also described as motivational orientation, e.g., Weingart, Bennet and Brett1993), a generalized goal that affects behavior in a given situation. Much of the research on goal orientation has taken place in educational settings, and has examined how student behavior is affected by a learning versus performance goal orientation (e.g., Dweck 1986). This notion has also been examined in an organizational setting, and is often focused on negotiation and conflict resolution (e.g., Button, Mathieu and Zajac 1996).

identity theory (Abrams and Hogg 1990) suggests that when an individual self-identifies as a member of a group, the unit of analysis for decision making shifts away from the individual and toward the group as a whole (Tajfel and Turner 1986). As a result, such employees are less likely to compete with one another, but may instead collude to maximize joint outcomes. Experimental evidence provides support for this notion, demonstrating that social identity results in higher levels of coordination (e.g., Wit and Wilke, 1992). Accounting policies can also affect mindset. For example, Rowe (2004) finds that the accounting report structure can create a "group frame" which increases cooperation among team members. As described earlier, cooperation in a tournament setting can take the form of collusion, and so by logical extension, the form of accounting reports could lead to a collusive mindset. Our intent here is not to provide a fully developed taxonomy of the various factors that influence mindset. Rather, we take it as given that employees may enter into a tournament with a general inclination toward either collusion or competition. In the next section, we describe how mutual monitoring amplifies the effect of mindset on effort and present our hypotheses.

Hypotheses

We predict that mutual monitoring will amplify the effect of mindset on effort. If employees enter into a tournament with collusive mindsets, they are likely to attempt collusion in order to lower effort. However, as described earlier, it is likely to be difficult to maintain collusive agreements without the ability to monitor the effort of other employees. Thus, we predict that for employees with collusive mindsets, mutual monitoring will facilitate collusion, leading to *lower* effort.

H2: When employees have collusive mindsets, effort will be lower when employees can mutually monitor one another than when they cannot.

In contrast, we predict that when employees enter the tournament with competitive mindsets, the effect of mutual monitoring will be *higher* effort. Recall the results of Study 1, in which very few collusive attempts were made and mutual monitoring led to increased effort. As described above, responses to a post-experimental question indicate that these participants had a strong desire to win the tournament, which suggests a competitive mindset. We conjecture that mutual monitoring led to increased effort in Study 1 because mutual monitoring increased the salience of the other tournament participants. Given the competitive mindset, this increased salience led participants to engage in social comparison, which led to increased effort. Thus, we predict that for employees with competitive mindsets, mutual monitoring will amplify the competitiveness, leading to higher effort.

> H3: When employees have competitive mindsets, effort will be higher when employees can mutually monitor one another than when they cannot.

In summary, our refined theory predicts that mutual monitoring will "turn up the volume" on whatever mindset with which participants enter the tournament, resulting in *lower* effort when employees have a collusive mindset and *higher* effort when employees have a competitive mindset.

Experimental Design

The second experiment uses a laboratory tournament with the same parameters and instructions as in the first. We use a 2 x 2 (mutual monitoring x mindset) between-subjects design. Mutual monitoring is manipulated in the same way as in the first study. Because our main interest is in the effect of mutual monitoring, and how this effect differs depending on whether or not participants are inclined to collude (i.e., whether they have a collusive mindset), we employ a strong manipulation to increase the likelihood that participants internalize the appropriate mindset. Specifically, mindset is manipulated by making salient, during the instruction stage, either the collusive or competitive strategy for participation in the tournament.

In the collusive mindset condition, participants are told the following:

> It is important to notice that the total partnership compensation is the same regardless of the effort levels chosen by the individual employees. That is, a total of 300 points will be paid as compensation (200 points to one of you and 100 points to the other) regardless of how low your total effort is. Therefore, you and your partner can earn higher points in total if you both put in low effort (because there is a cost to choosing a higher effort level).

On the other hand, in the competitive mindset condition, participants are told the following:

> It is important to notice that your compensation depends on whether your effort is high enough to make your performance better than the other employee's performance. That is, your compensation will be 100 points higher as long as your performance is better than the other employee's performance. Therefore, you will receive the higher compensation level (200 points) if you choose high enough effort to beat the other employee's performance, even if only by a little bit.

All other aspects of the laboratory tournament are identical to Study 1.[15]

Participants and Procedures

Participants are undergraduate business students from the same population used for Study 1.[16] Four experimental sessions were conducted, one for each experimental cell. Three of the sessions had ten participants and the fourth (mutual monitoring/collusive mindset) had eight participants.

Results

In this section, we provide formal tests of the hypotheses and additional analyses. Descriptive statistics related to the dependent variable and post-experimental questions are reported in Table 3.

---

[15] The one exception relates to the descriptions of payoff structures in the two studies. In Study 1, the payoffs were described as having both employees receive base pay, with the winner receiving a bonus. In Study 2, the description did not separate the payoffs into base pay and bonus. Rather, participants were told the total payoffs that would be received by the participants with the higher and lower performance. We made this change to eliminate any framing effects that may interfere with the mindset manipulation.

[16] The average age of participants is 18.4 years. They report 0.7 years of work experience, on average. Seventy-one percent of the participants are male.

Hypotheses Tests

H2 predicts that when employees are in a collusive mindset, mutual monitoring will facilitate collusion, thus leading to lower effort. H3, on the other hand, predicts that when employees are in a competitive mindset, mutual monitoring will reinforce competitiveness via social comparison process, thus leading to higher effort. Together, H2 and H3 predict an interaction, such that the effect of mutual monitoring on effort will depend on the mindset of employees.

As before, our dependent measure is the average effort for the two participants in each unique dyad across the four periods in which the dyad interacted. The pattern of results is presented in Figure 2. As reported in Table 4, Panel A, a two-way ANOVA with effort as the dependent variable, finds a significant two-way interaction between mutual monitoring and mindset ($F = 21.03$, $p < 0.01$, two-tailed). A simple effects test (Table 4, Panel B) supports H2. Specifically, in the collusive mindset condition, mean effort is lower with mutual monitoring (4.93) compared to without (14.88) and this difference is statistically significant ($F = 7.47$, $p < 0.01$, two-tailed). The simple effects test does not support H3, however, because in the competitive mindset condition, mutual monitoring does not have a significant effect on effort ($F = 0.20$, $p = 0.66$, two-tailed).[17]

We are able to shed light on the initial lack of support for H3 by examining period-by-period results. Because each unique dyad interacts for four periods, we collect four within-subjects data points (periods) for each dyad (see Table 3). Using only those dyads in the competitive mindset condition, we conduct a repeated measures ANOVA, in which period serves

[17] For robustness, we also conduct ANOVAs for each cycle, using mean dyad effort for the cycle as the dependent variable. Recall that three sessions had five dyads per cycle and one session had four dyads per cycle. As such, this test has four or five observations per cell. Despite this low power, we find a significant interaction between mutual monitoring and mindset for each of the four cycles, with two-tailed p-values of 0.03, 0.002, 0.07 and 0.09 for cycles one through four, respectively.

as the within-subjects independent variable, mutual monitoring serves as the between-subjects dependent variable, and average dyad effort serves as the dependent measure. We find a significant interaction between period and mutual monitoring ($F = 2.89$, $p = 0.05$, two-tailed, untabulated), indicating that the effect of mutual monitoring differs across periods. Further, using the method of orthogonal polynomial contrasts, we compare linear trends across conditions and find that the trend in effort is more positive with mutual monitoring than without ($F = 1.66$, $p = 0.10$, one-tailed, untabulated). In fact, by Period 4, effort is higher with mutual monitoring than without (15.78 vs. 20.30, $p = 0.03$, one-tailed, untabulated). These results provide evidence of the potential for mutual monitoring to increase effort when participants are inclined to compete, consistent with the intuition behind H3.[18]

We do acknowledge, however, that the positive effect of mutual monitoring on effort is not nearly as strong as observed in Study 1, and we can only conjecture as to why this is the case. As noted in Footnote 15, the only difference in our experimental procedures for the two studies (other than the manipulation of mindset) is that in Study 1, the tournament was described using a bonus frame, whereas in Study 2, the framing was neutral. It is possible that the bonus frame created an even more competitive mindset than the language used to invoke a competitive mindset in Study 2. In the next section, we report additional analysis supporting our conjecture that mutual monitoring "turns up the volume" on competitiveness. It, therefore, seems likely that mutual monitoring could affect effort more dramatically in a situation in which the competitive mindset was quite strong, as we observe in Study 1.

Additional Analyses

---

[18] It is possible that participants exerted additional effort in period 4 because they would not need to fear that their own higher effort would lead to a ratcheting up of effort in future periods. Although the data from Study 2 do not allow us to exclude this possibility, the results of Study 1 provide some confidence that this result is not obtained solely due to last period effects.

In this section, we report additional analyses in order to 1) gain greater insight into the processes responsible for the effects we observe and 2) provide a more detailed understanding of how the behavior we observe compares to the basic economic predictions in the strategic context we model and examine.

We use responses to the post-experimental questionnaire to gain a deeper understanding of the process, resulting in three key insights. The first insight is that, consistent with our mindset manipulation, we find evidence that participants in the collusive mindset condition were more likely to attempt collusion than participants in the competitive mindset condition. Two pieces of evidence support this conclusion. First, a question in the post-experimental questionnaire (Table 3, Question 9) asks participants whether they ever attempted to come to an agreement with the other participant on what their effort levels would be. While only 30 percent (6/20) of participants in the competitive mindset condition report such an attempt, 100 percent of participants in the collusive mindset condition report they attempted to do so. This (untabulated) difference is statistically significant ($\chi^2 = 19.95$, $p < 0.01$, two-tailed). Second, two independent coders (who were unaware of the hypotheses or manipulations) examined the text from the chat sessions between each dyad to identify whether or not each dyad agreed to collude.[19] This process determined that 25 percent of dyads agreed to collude in the competitive mindset condition, whereas 94 percent agreed in the collusive mindset condition. This (untabulated) difference is also statistically significant ($\chi 2 = 37.48$, $p < 0.01$, two-tailed).[20]

The second insight is that, consistent with our conjecture from Study 1, mutual monitoring "turns up the volume" on the competitive mindset. We obtain this insight by

[19] Cohen's Kappa for inter-rater reliability is 0.919 (p<0.001). Coder disagreements were resolved by one of the co-authors.
[20] If we conduct a binary logistic regression, including both mutual monitoring and the mutual monitoring X mindset interaction, neither the main effect for mutual monitoring nor the interaction is significant.

conducting a factor analysis using four questions (Table 3, Questions 1 through 4) intended to capture the participant's relative focus on competitive (Questions 1 and 2) versus collusive (Questions 3 and 4) strategies. The factor analysis obtains two factors with eigenvalues greater than 1.00. The first, with an eigenvalue of 1.94, loads as would be expected, with positive loadings for the first and second questions and negative loadings for the third and fourth questions. We use this factor score as our measure of the participant's focus on competitive versus collusive strategies.[21] We would naturally expect the mindset manipulation to affect this factor score, with participants in the competitive mindset condition scoring higher than those in the collusive mindset condition. Importantly, when we conduct an (untabulated) ANOVA on the factor score, we find an ordinal interaction between mutual monitoring and mindset ($F = 4.34$, $p = 0.05$, two-tailed). Simple effects analysis suggests that mutual monitoring magnifies the competitiveness mindset in the competitive condition ($F = 3.04$, $p = 0.10$, two-tailed), but has no effect on mindset in the collusive condition ($F = 1.48$, $p = 0.23$, two-tailed). This finding is consistent with our conjecture made from Study 1 that when employees are in a competitive mindset, mutual monitoring increases their competitiveness, likely by increasing the salience of social comparison.

Our third insight relates to the underlying process between mindset and usage of information obtained from mutual monitoring. Our theory suggests that the use of such information will depend on the mindset with which participants enter the tournaments. Responses to the post-experimental questionnaire allow us to gain insight into whether such differences in informational usage are responsible for the effects we observe when testing our hypotheses. Participants in the mutual monitoring condition were asked how they used the

[21] The second score, with an eigenvalue of 1.12 loads positively on questions 2 and 3 and negatively on questions 1 and 4. It is unclear what underlying construct is captured by this factor score, and we ignore it in our analysis.

information about the other participant's effort choices (Table 3, Questions 5 through 8). The first two of these questions indicate the use of mutual monitoring for competitive purposes, whereas the latter two indicate its use for collusive purposes. We conduct a factor analysis, obtaining two factors with eigenvalues greater than 1.00. The first, with an eigenvalue of 1.61, loads as would be expected, with positive loadings for the first and second questions and negative loadings for the third and fourth questions. We use this factor score as our measure of the participant's use of information for competitive versus collusive purposes.[22] We then conduct an ANOVA, and the results are consistent with the processes suggested by our theory. Specifically, a significant main effect for mindset ($F = 10.51$, $p < 0.01$, two-tailed) confirms that participants in the competitive mindset condition are more likely to use the information gathered through mutual monitoring for competitive purposes. This analysis supports our underlying theory that the mindset with which participants enter the tournament affects whether information from mutual monitoring will be used for competitive or collusive purposes.

Finally, we compare the results in all four experimental conditions to the benchmarks provided by the game theoretic analysis of the tournament setting. Recall that the Nash equilibrium has both employees choosing effort levels of 17.5, whereas the Pareto optimal outcome has both choosing effort levels of 1.0. Also recall that Table 3 reports and Figure 2 presents the pattern of results for our main dependent variable, mean effort. As reported in Table 5, additional analysis finds that participants in the competitive mindset condition chose effort at a level that is statistically indistinguishable from the Nash equilibrium of 17.5, regardless of whether or not mutual monitoring is possible (mutual monitoring: $p = 0.49$, two-tailed; no

[22] The second score, with an eigenvalue of 1.22 loads positively on all four questions. It is unclear what underlying construct is captured by this factor score, and we ignore it in our analysis.

mutual monitoring: $p = 0.61$, two-tailed).[23] However, in the collusive mindset condition, effort levels are statistically distinguishable from the Nash equilibrium of 17.5 both when mutual monitoring is possible ($p < 0.01$ two-tailed) and when it is not ($p = 0.06$ two-tailed). Further, participants in all four conditions provide effort at levels greater than the Pareto optimal level of 1.0 (all $p < 0.01$). In summary, although mean effort in the collusive/mutual monitoring condition is very low (4.93), participants were unable to achieve *perfect* collusion.

Next, we investigate whether the actual outcomes are closer to the Nash equilibrium or the Pareto optimal solution. Specifically, for each condition we compare the distance between the observed mean effort and the Nash equilibrium to the distance between the observed mean effort and the Pareto optimal outcome. We find that in the competitive mindset condition, mean effort is significantly closer to the Nash equilibrium than the Pareto optimal outcome regardless of whether or not mutual monitoring is possible (both $p < 0.01$). In the collusive mindset condition, however, results depend on whether mutual monitoring is possible. With mutual monitoring, mean effort is significantly closer to the Pareto optimal outcome than the Nash equilibrium ($p < 0.01$). Without mutual monitoring, however, mean effort level is significantly closer to the Nash equilibrium ($p < 0.01$). These results suggest that although the Nash equilibrium is highly descriptive of behavior in most conditions, the combination of a collusive mindset and the ability to mutually monitor leads to more successful collusion, and thus lower effort, in our tournament setting.

The fact that mutual monitoring has such a strong impact on whether participants in the collusive mindset condition tend towards the Nash equilibrium versus Pareto optimal outcomes is particularly striking given that 100 percent of these participants report having attempted to

[23] In fact, when mutual monitoring is possible, participants with competitive mindsets actually put in greater effort than the Nash equilibrium level in Period 4 ($t = 2.44$, $p = 0.02$ two-tailed), perhaps because participants have a strong desire to win and mutual monitoring reinforces competition.

collude by agreeing with the other participants on effort levels. Despite such attempts, when mutual monitoring is not possible, these participants' effort levels are more closely represented by the (non-cooperative) Nash equilibrium than by the (cooperative) Pareto optimal outcome. These results speak to the inherent instability of collusion in this setting and the potential for mutual monitoring to be used to enforce collusive agreements, leading to a reduction in effort to the detriment of the firm.

IV: DISCUSSION

This study investigates how mutual monitoring affects effort when individuals are compensated via rank-order tournaments. The prior literature generally assumes that mutual monitoring will have a detrimental effect on effort in tournaments, because it will facilitate collusion among tournament participants. Our two experimental studies show that this is not necessarily the case. In our first study, we find that mutual monitoring actually increases effort, because participants do not attempt to collude but rather behave competitively. The results of Study 1 lead to a more refined theory regarding the role of mutual monitoring in tournaments in which we introduce the concept of mindset. Our refined theory predicts that the effect of mutual monitoring on effort will depend on the mindset with which employees enter the tournament. Specifically, we predict that mutual monitoring will "turn up the volume" on whatever mindset participants have when they enter the tournament, resulting in *lower* effort when employees have a collusive mindset and *higher* effort when employees have a competitive mindset.

Results of our second study support this refined theory. Specifically, we find that mutual monitoring *decreases* effort when participants enter the tournament with collusive mindsets. In contrast, by the final period of interaction, mutual monitoring *increases* effort when participants enter the tournament with competitive mindsets. Further analysis of the data suggests that the process through which mutual monitoring “turns up the volume” depends on the mindset. Specifically, when participants have a collusive mindset, mutual monitoring facilitates collusion by providing information useful for enforcing collusive agreements. Without mutual monitoring, however, attempts to collude unravel such that effort is significantly closer to the competitive Nash equilibrium than the Pareto optimal collusive outcome. When participants have a competitive mindset, mutual monitoring leads to increased effort by reinforcing competitiveness. Apparently, mutual monitoring makes other participants more salient, which activates a social comparison process through which competitiveness becomes magnified.

Our experimental evidence provides a potential explanation for why tournament incentive schemes are frequently observed in the work place, despite the theoretical threat of collusion. Specifically, we show that collusion is unlikely to be a serious threat to effort unless employees have collusive mindsets and the ability to mutually monitor one another. That is, both of these conditions are necessary for successful collusion, but neither is sufficient. In fact, we find evidence that in some circumstances, mutual monitoring can actually increase effort under a tournament contract. This finding is a contradiction to the maintained assumption that mutual monitoring will be detrimental to effort in tournaments.

Importantly, both the degree of mutual monitoring and employee mindset can be directly influenced by managerial decisions. Our study suggests that it is important for managers to evaluate whether such decisions will have positive or negative effects when considered in conjunction with incentive compensation schemes. For example, many firms that advocate the use of the forced ranking system invest great energy and money into building social ties among employees because it is believed that the cooperation induced by a strong team spirit will translate into increased performance.[24] Our results, however, suggest that if this cooperative spirit leads to collusion among employees, then it may have a negative effect on effort, depending on other incentive arrangements that are in place. Consistent with this view, Guralnik, Rozmarin and So (2004) observe that promoting teamwork and simultaneously establishing a forced ranking system can potentially send a double message to the employees. They suggest that in order for team building to provide positive effects in firms with a forced ranking system, managers should attempt to include team building as *part of the competition*.

More generally, our results imply that various management control devices must be considered as parts of an overall system of control. For example, both experimental (Zhang 2008) and field (Chen and Sandino 2010) studies suggest that management practices that are perceived as unfair can lead employees to develop a collusive mindset. Our study suggests that this effect will be particularly problematic in firms that use forced ranking or other tournament compensation schemes in environments where mutual monitoring is possible. Interactive effects of various control mechanisms, such as these, are likely highly complex. They become even more so when one begins to incorporate the effects of informal controls (e.g., trust) as well. This observation suggests that examinations of the rich and complex interrelations among both formal and informal controls offer significant future research opportunities.

[24] Consider General Electric, for example. Their forced ranking system explicitly states that an employee's ability to energize those around him/her is an important criterion for ranking (Grote, 2005). Similarly, Exxon Mobil is known for its forced ranking system, and yet a number of consultants list this firm as a client of their teambuilding practices (e.g., First Events, http://www.feetfirstevents.com/team-building/corporate-teambuilding-clients.html; Afterburner: http://www.afterburnerseminars.com/leadership-team-building; Quixote Consulting: http://www.quixoteconsulting.com/aboutus.html).

REFERENCES

Abrams, D., and M. A. Hogg. 1990. An introduction to the social identity approach. In *Social Identity Theory: Constructive and Critical Advances*, edited by D. Abrams and M. A. Hogg. New York: Springer-Verlag New York, Inc.

Arya, A., J. C. Fellingham, and J. C. Glover. 1997. Teams, repeated tasks, and implicit incentives. *Journal of Accounting and Economics* 23 (1): 7-30.

Baiman, S. 1982. Agency research in managerial accounting: a survey. *Journal of Accounting Research* 1: 154-213.

Bandiera, O., I. Barankay, and I. Rasul. 2005. Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics* 120 (3): 917–962.

Bogaert, S., C. Boone and C. Declerck. 2008. Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology* 47(3): 453-480.

Boyle, M. 2001. Performance Reviews: Perilous Curves Ahead. *Fortune*, 15 May 2001.

Bull, C., A. Schotter, and K. Weigelt. 1987. Tournaments and piece rates: an experimental study. *The Journal of Political Economy* 95 (1): 1-33.

Button, S. B., J. E. Mathieu, and D. M. Zajac. 1996. Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes* 67(1): 26-48.

Carpenter J., S. Bowles, and H. Gintis. 2006. Mutual monitoring in teams: Theory and experimental evidence on the importance of reciprocity. *IZA Discussion Paper* No. 2106.

Chen, C. X., and T. Sandino. 2010. Can wages buy honesty? The relationship between relative wages and employee theft. Working paper, University of Illinois at Urbana-Champaign and University of Southern California.

Cooper, R., D. V. DeJong, R. Forsythe and T. W. Ross. 1996. Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior* 12: 187-218.

De Dreu, C. K. W., B. Beersma, K. Stroebe, and M. C. Euwema. 2006. Motivated information processing, strategic choice, and the quality of negotiated agreement. *Journal of Personality and Social Psychology* 90(6): 927-943.

Dweck, C. S. 1986. Motivational processes affecting learning. *American Psychologist* 41(10): 1040-1048.

Dye, R. 1984. The trouble with tournaments. *Economic Inquiry* 22 (1): 147-149.

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

Festinger, L. 1954. A theory of social comparison processes. *Human Relations* 7 (2): 117-140.

Fischbacher, U. 2007. Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics* 10(2): 171-178.

Förster, J., N. Liberman, and R. S. Friedman. 2009. What do we prime? On distinguishing between semantic priming, procedural priming, and goal priming. In *Oxford Handbook of Human Action*, edited by E. Morsella, J. A. Bargh, and P. M. Gollwitzer, 173-196. New York, NY: Oxford University Press, Inc.

Frederickson, J. R. 1992. Relative performance information: the effects of common uncertainty and contract type on agent effort. *The Accounting Review* 67 (4): 647-669.

Frederickson, J. R. and W. Waller. 2005. Carrot or stick? Contract framing and the use of decision-influencing information in a principal–agent setting. *Journal of Accounting Research* 43(5): 709-733.

Gibbons, R. and K. J. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial and Labor Relations Review* 43 (3): 30S-51S.

Gollwitzer, P. M., and U. Bayer. 1999. Deliberate versus implemental mindsets in the control of action. In *Dual Process Theory in Social Psychology*, edited by S. Chaiken, and Y. Trope, 403-423. New York, NY: The Guilford Press.

Green, E. J. and R. H. Porter. 1984. Non-cooperative collusion under imperfect price information. *Econometrica* 52 (1): 87-100

Green, J. R. and N. L. Stokey. 1983. A Comparison of Tournaments and Contracts. *The Journal of Political Economy* 91(3): 349-364.

Grote, R. C. 2005. Forced ranking: Making performance management work. *Harvard Business School Working Knowledge*, 14, November.

Guralnik. O, E. Rozmarin and A. So. 2004. Forced distribution: is it right for you? *Human Resource Development Quarterly* 15 (3): 339-345.

Hannan, R. L. 2005. The combined effect of wages and firm profit on employee effort. *The Accounting Review.* 80(1):167-189.

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.

Harbring, C. 2006. The effect of communication in incentive systems - An experimental study. *Managerial and Decision Economics* 27: 333-353.

Harinck, F. and C. K. W. De Dreu. 2008. Take a break! Or not? The impact of mindsets during breaks on negotiation processes and outcomes. *Journal of Experimental Social Psychology* 44: 297-404.

Hofstede, G. 1980. *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA: Sage Publications.

Hofstede, G. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations.* Thousand Oaks CA: Sage Publications.

Holmstrom B. and P. Milgrom. 1990. Regulating trade among agents. *Journal of Institutional and Theoretical Economics* 146: 85–105.

Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole. 2003. The Economics of tacit collusion. *Final Report for DG Competition, European Commission*.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson. 1982. Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory* 27: 245-252.

Ku, G., D. Malhotra and J. K. Murnighan. 2005. Towards a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions. *Organizational Behavior and Human Decision Processes* 96: 89–103.

Kuang, X. J., and D. V. Moser. 2009. Reciprocity and the effectiveness of optimal agency contracts. *The Accounting Review* 84(5): 1671-1694.

Kühn, K. 2001. Fighting collusion by regulating communication between firms. *Economic Policy* 16 (32): 167-204.

Lambert, R. A., D. F. Larcker, and K. Weigelt. 1993. The structure of organizational incentives. *Administrative Science Quarterly* 38(3): 663-691.

Lazear, E, P. 2000. Performance pay and productivity. *American Economic Review* 90(5): 1346-1361.

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

Ma, C. 1988. Unique implementation of incentive contracts with many agents. *Review of Economic Studies* 55 (4): 555-572.

Mas, A. and E. Moretti. 2009. Peers at work. *American Economic Review* 99(1): 112-145.

McGregor, J. 2006. The struggle to measure performance. *Business Week* January 9, 2006.

Motta, M. 2004. *Competition Policy- Theory and Practice*. Cambridge University Press. Cambridge, UK.

Nalebuff, B. J., and J. E. Stiglitz. 1983. Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics* 14 (1): 21-43.

O'Connell, V. 2008. Retailers reprogram workers in efficiency push. *Wall Street Journal* (September 10, 2008), A1.

Orrison, A., A. Schotter and K. Weigelt. 2004. Multi person tournaments: an experimental examination. *Management Science* 50(2):268-279.

Potters, J., B. Rockenbach, A. Sadrieh and E. van Damme. 2004. Collusion under yardstick competition: An experimental study. *International Journal of Industrial Organization* 22(7): 1017-1038.

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

Rowe, C. 2004. The effect of accounting report structure and team structure on performance in cross-functional teams. *The Accounting Review* 79(4): 1153-1180.

Selten, R. and R. Stoecker. 1986. End behavior in sequences of finite prisoner's dilemma supergames. *Journal of Economic Behavior and Organization* 7: 47-70.

Stigler, G. J. 1964. A Theory of Oligopoly. *Journal of Political Economy* 72 (1): 44-61.

Tafkov, I. 2010. Private and public relative performance information under different incentive systems. Working Paper, Georgia State University.

Tajfel, H., and J. Turner. 1986. The social identity theory of intergroup behavior. In *The Social Psychology of Intergroup Relations (2nd ed.),* edited by S. Worchel and W. G. Austin. Chicago: Nelson-Hill Publishers.

Tesser, A. and J. Campbell. 1980. Self-definition: The impact of the relative performance and similarity of others. *Social Psychology Quarterly* 43(3): 341–347.

Tirole, J. 1988. *The Theory of Industrial Organization*. M.I.T. Press. Cambridge, MA.

Towry, K. L. 2003. Control in a teamwork environment—the impact of social ties on the effectiveness of mutual monitoring contracts. *The Accounting Review* 78 (4): 1069–1095.

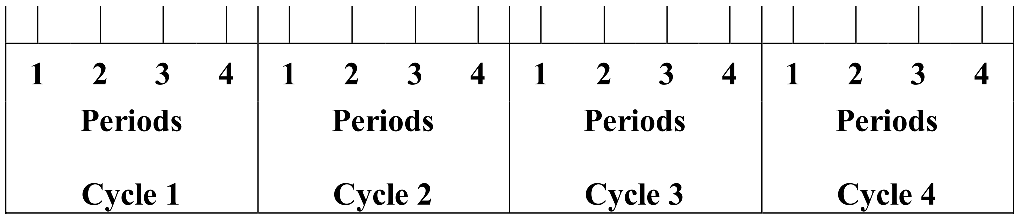
Weingart, L. R., R. J. Bennet, and J. M. Brett. 1993. The impact of consideration of issues and motivational orientation on group negotiation process and outcome. *Journal of Applied Psychology* 78(3): 504-517.

Wit, A. P., and H. A. M. Wilke. 1992. The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology* 13(1): 135-151.

Zhang, Y. 2008. The effects of perceived fairness and communication on honesty and collusion in a multi-agent setting. *The Accounting Review* 83 (4): 1125-1146.

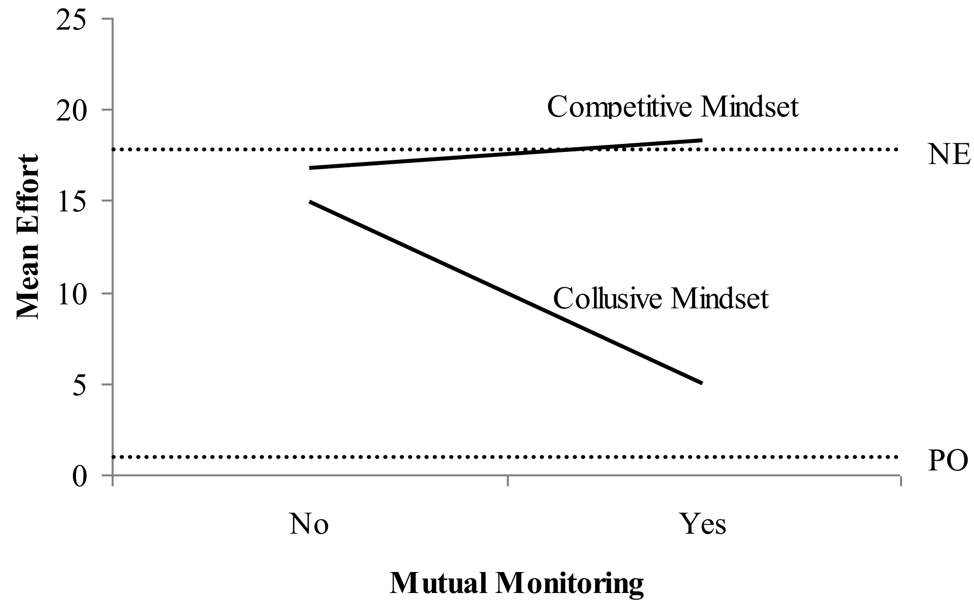
FIGURE 1

Experimental Timeline



Dyads are matched at the beginning of each cycle, then interact for four periods. A new cycle begins after the end of the 4th period, and dyads are re-matched. This process is repeated for a total of four cycles in each experimental session.

FIGURE 2
The Effects of Mutual Monitoring and Mindset on Effort



Mindset: Participants in the "Competitive Mindset" condition receive instructions that emphasize the benefit of outperforming the other employee. Participants in the "Collusive Mindset" condition receive instructions that emphasize the benefit of saving the cost of effort.

Mutual Monitoring: Participants are able to observe (not observe) the other employee's effort level when there is (is not) mutual monitoring.

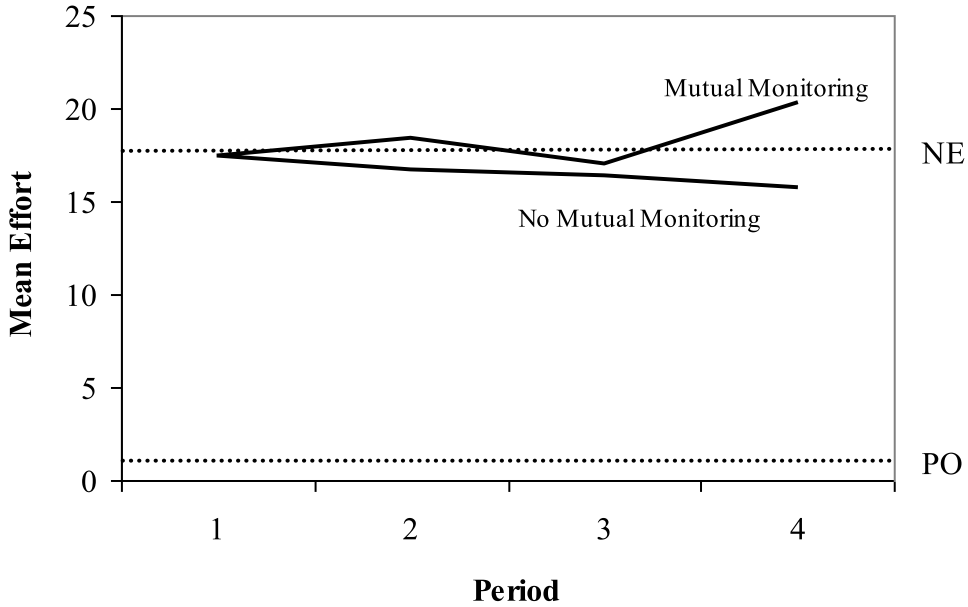
Effort: Effort is chosen by participants, with a range of between 1 and 40. The cost of effort increases in the level chosen and expected performance increases in effort.

NE: Nash Equilibrium effort level (17.5)

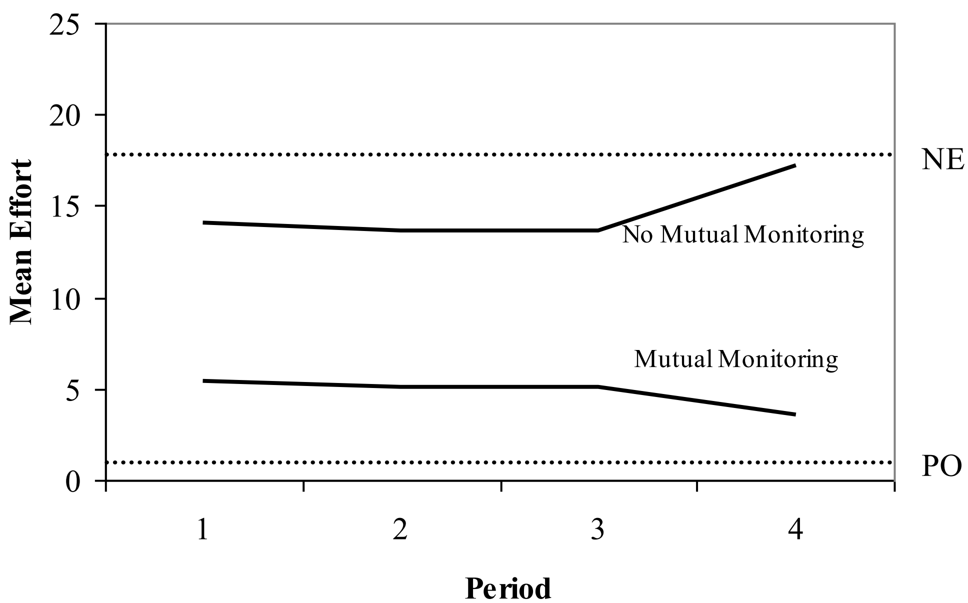
PO: Pareto Optimal effort level (1.0)

Figure 3
Period Trends

Panel A: Competitive Mindset



Panel B: Collusive Mindset



See Figure 2 for a description of the variables

TABLE 1
Cost of Effort[a]

Effort	Cost	Effort	Cost	Effort	Cost	Effort	Cost
1	0.1 point	11	17.3 points	21	63.0 points	31	137.3 points
2	0.6 point	12	20.6 points	22	69.1 points	32	146.3 points
3	1.3 point	13	24.1 points	23	75.6 points	33	155.6 points
4	2.3 points	14	28.0 points	24	82.3 points	34	165.1 points
5	3.6 points	15	32.1 points	25	89.3 points	35	175.0 points
6	5.1 points	16	36.6 points	26	96.6 points	36	185.1 points
7	7.0 points	17	41.3 points	27	104.1 points	37	195.6 points
8	9.1 points	18	46.3 points	28	112.0 points	38	206.3 points
9	11.6 points	19	51.6 points	29	120.1 points	39	217.3 points
10	14.3 points	20	57.1 points	30	128.6 points	40	228.6 points

[a]Effort is chosen by participants, with a range of between 1 and 40. The cost of effort increases in the level chosen and expected performance increases in effort.

Participants were provided with a similar table during the experiment.

Table 2
Study 1 Descriptive Statistics

	Mutual Monitoring [a]	
Mean (Std. Dev.)	No	Yes
Period 1 Effort [b]	18.23 (4.19)	21.69 (7.31)
Period 2 Effort	21.69 (7.31)	23.75 (6.23)
Period 3 Effort	17.85 (7.91)	22.88 (8.67)
Period 4 Effort	18.03 (6.35)	22.95 (5.80)
Total Effort	18.73 (3.81)	22.95 (5.80)
Collusive Attempts [c]	0% (= 0/8)	30% (= 3/10)

Notes:

[a] *Mutual Monitoring*: Participants are able to observe (or not observe) the other employee's effort level when there is (or is not) mutual monitoring.

[b] *Effort*: Effort is selected by participants, with a range of between 1 and 40. The cost of effort increases in the level chosen and expected performance increases in effort.

[c] *Collusive Attempts*: Participants were asked whether they ever attempted to come to an agreement with the other participant on what their effort levels would be. Reported values are the proportion of participants who answered "Yes."

Table 3
Study 2 Descriptive Statistics

Panel A: Organized by Mindset Condition

Mean (Std. Dev.)	Mindset [a]			
	Competitive		Collusive	
	Mutual Monitoring [b]		Mutual Monitoring	
	No	Yes	No	Yes
Period 1 Effort [c]	17.40 (6.25)	17.45 (6.45)	14.03 (7.56)	5.38 (7.21)
Period 2 Effort	17.10 (6.74)	18.38 (7.45)	13.65 (7.43)	5.09 (6.17)
Period 3 Effort	16.82 (8.00)	17.05 (7.02)	14.60 (7.28)	5.69 (7.28)
Period 4 Effort	15.78 (9.40)	20.30 (5.11)	17.25 (7.34)	3.56 (3.05)
Total Effort	16.78 (6.20)	18.29 (4.99)	14.88 (5.92)	4.93 (4.09)
Q1:Choose higher effort [d]	4.00 (1.56)	5.10 (0.57)	3.30 (1.34)	3.00 (2.14)
Q2:Perform better than the other employee	5.60 (0.84)	5.80 (1.03)	5.80 (1.55)	5.13 (1.13)
Q3:Cooperate with the other employee	4.70 (1.34)	2.80 (1.23)	4.60 (2.67)	5.75 (1.49)
Q4:Keep effort low	4.90 (1.60)	5.10 (1.66)	6.10 (0.99)	6.25 (1.04)
Q5:Use information to outperform the other employee [e]	N/A	4.60 (2.01)	N/A	4.13 (1.46)
Q6:Use information to learn effort required to win	N/A	6.00 (0.67)	N/A	4.38 (1.41)
Q7:Use information to ascertain low effort	N/A	5.50 (1.08)	N/A	5.88 (1.13)

Table 3 (Continued)

	Mindset			
	Competitive		Collusive	
	Mutual Monitoring		Mutual Monitoring	
Mean (Std. Dev.)	No	Yes	No	Yes
Q8:Use information to learn about reneging	N/A	3.60 (1.65)	N/A	5.38 (2.00)
Q9:Try to agree on an effort level [f]	50% (=5/10)	10% (=1/10)	100% (=10/10)	100% (=8/8)
Q10:Tempted to violate the agreement	3.80 (2.39)	1.00 (N/A)	5.00 (2.11)	4.13 (2.30)
Q11:Nervous to be betrayed	4.40 (2.07)	1.00 (N/A)	5.50 (1.78)	4.88 (1.25)
Collusive Attempts [g]	40% (= 8/20)	10% (= 2/20)	90% (= 18/20)	100% (=16/16)
Number of Unique Dyads	20	20	20	16
Number of Participants	10	10	10	8

Table 3 (Continued)

Notes:
[a] *Mindset:* Participants in the "Competitive Mindset" condition receive instructions that emphasize the benefit of outperforming the other employee. Participants in the "Collusive Mindset" condition receive instructions that emphasize the benefit of saving the cost of effort.

[b] *Mutual Monitoring*: Participants are able to observe (or not observe) the other employee's effort level when there is (or is no) mutual monitoring.

[c] *Effort*: Effort is selected by participants, with a range of between 1 and 40. The cost of effort increases in the level chosen and expected performance increases in effort.

[d] *Q1 through Q11* are responses to questions from the post-experimental questionnaire.

Q1 through Q4 asked participants to indicate, on a seven-point Likert scale (from 1= "very unimportant" to 7="very important"), how important it had been to them when choosing an effort level to
Q1: choose a higher effort level than the other employee
Q2: perform better than the other employee
Q3: cooperate with the other employee
Q4: keep their effort level low

[e] Q5 through Q8 asked participants to indicate, on a seven-point Likert scale (from 1 = "not at all" to 7 = "very much"), to what degree they used the information to
Q5: try to make their own effort level higher than the other employee's effort level
Q6: learn the effort level required to obtain higher performance than the other employee
Q7: ascertain whether the other employee chose a low effort level so that the respondent could choose a low effort level too
Q8: learn if the other employee reneged on an agreed effort level.

[f] Q9 asked participants whether they ever attempted to come to an agreement with the other participant on what their effort levels would be. Reported values are the proportion of participants who answered "Yes."

If the answer to Q9 was "Yes," participants were asked Q10 and Q11.
Q10 asked the participants to indicate, on a seven-point Likert scale (from 1 = "not at all tempted" to 7 = "very tempted"), how tempted they were to violate their agreement.
Q11asked the participants to indicate, on a seven-point Likert scale (from 1 = "not at all nervous" to 7 = "very nervous"), how nervous they were that their partner would violate the agreement.

[g] *Collusive Attempts:* Based on coding of chat transcripts by independent coders.

TABLE 4
The Effects of Mutual Monitoring and Mindset on Effort [a]

Panel A: ANOVA results

Factor	*dof*	Mean Square	*F*	*p*-value (two-tailed)
Mindset	1	4,382.13	37.21	<0.01
Mutual Monitoring	1	1,338.59	11.37	<0.01
Mindset x Mutual Monitoring	1	2,476.58	21.03	<0.01
Error	72	117.78		
Panel B: Simple Effects for each Mindset Condition				
Effect of Mutual Monitoring: *Collusive* Mindset	1	880.30	7.47	<0.01
Effect of Mutual Monitoring: *Competitive* Mindset	1	23.07	0.20	0.66
Panel C: Planned Comparison, Mutual Monitoring Condition Only				
Effect of Mindset	1	1,587.54	13.48	<0.01

Notes:
[a] See Table 3 for a description of variables.

TABLE 5
Comparison of Effort Levels to Economic Benchmarks

	Mindset			
	Competitive Mutual Monitoring		Collusive Mutual Monitoring	
	No	Yes	No	Yes
Effort $_{\textbf{actual}}$ [a]	16.78	18.29	14.88	4.93
Effort $_{\textbf{NE}}$ [b]	17.50	17.50	17.50	17.50
Effort $_{\textbf{PO}}$ [c]	1.0	1.0	1.0	1.0
Effort $_{\textbf{actual}}$ **vs. Effort** $_{\textbf{NE}}$	t=-0.52 p=0.61 (two-tailed)	t=0.71 p=0.49 (two-tailed)	t=-1.98 p=0.06 (two-tailed)	t=-12.28 p<0.01 (two-tailed)
Effort $_{\textbf{actual}}$ **vs. Effort** $_{\textbf{PO}}$	t=12.10 p<0.01 (two-tailed)	t=16.40 p<0.01 (two-tailed)	t=11.25 p<0.01 (two-tailed)	t=4.82 p<0.01 (two-tailed)
|Effort $_{\textbf{actual}}$ **-. Effort** $_{\textbf{NE}}$**| vs. |Effort** $_{\textbf{actual}}$ **-. Effort** $_{\textbf{PO}}$**|**	t=5.71 p<0.01 (two-tailed)	t=10.00 p<0.01 (two-tailed)	t=4.40 p<0.01 (two-tailed)	t=4.22 p<0.01 (two-tailed)

Notes:
[a] *Effort* $_{actual}$: study participants' mean effort level observed in each of the four experimental conditions.

[b] *Effort* $_{NE}$: the benchmark Nash equilibrium effort level in each of the four experimental conditions.

[c] *Effort* $_{PO}$: the Pareto Optimal effort level in each of the four experimental conditions.