Comparing the strengths of self-interest and prosocial motivations
Ye Li and Margaret S. Lee

Abstract
As organizations move towards more group-based work environments, individuals must make more decisions on how much effort they are willing to exert for the benefit of other people. Working for the benefit others inevitably takes away from the resources a person could invest towards his or her own benefit. This paper therefore tries to answer two questions: In what circumstances are people willing to work for the benefit others and how much effort are they willing to exert? Past research on prosocial behavior has not compared people’s willingness to exert effort for others versus for themselves. The goal of this research was to fill this gap by directly comparing the motivating strength of self-interest to that of prosocial motivations. In a series of studies, participants either worked for their own pay or for the payment of another participant (and no benefit to themselves). We found that under conditions in which a level of acceptable effort was not defined and a social cost would be incurred by not helping the other person, participants working on behalf of others exerted more effort (earning another participant more money) compared to participants working for their own pay. We discuss implications for group and workplace incentives.

KEYWORDS: worker motivation, incentives, prosocial behavior, impression management, organizational behavior
INTRODUCTION

Every action and decision has some underlying motivation. The motivation may be as simple as a craving for a certain food when choosing what to eat, or it can be as complex as a higher-level desire to maintain one’s moral values. The topic of motivation has interested researchers in many fields, including psychology (e.g., Maslow 1943, Deci 1971, Elliott and Dweck 1988, Baumeister and Leary 1995, Ryan and Deci 2000), economics (e.g., Kreps 1997, Prendergast 1999, Gneezy and Rustichini 2000, Frey and Jegen 2001, Fehr and Falk 2002, Benabou and Tirole 2003, 2006), and organizational studies (e.g., Locke 1968, Hackman and Oldham 1976, O'reilly et al. 1991). Motivation refers to the psychological processes that direct, energize, and sustain action and is usually elicited through the use of incentives (for review, see Latham and Pinder 2005). In the realm of work motivation, previous research has focused on extrinsic motivators, such as monetary compensation or employee benefits, and intrinsic motivators, such as job satisfaction and positive interpersonal relationships with coworkers. However, an area that needs more exploration is the motivating effect of other people. Barack Obama said to the graduating class of 2009 at Arizona State University: “Find somebody to be successful for. Raise their hopes. Think of their needs.” We seem to believe that others can inspire us to be greater, to achieve more, and to find greater meaning in our lives. We tend to think that parents are motivated to work harder when their efforts directly impact their children; dedicated teachers go beyond the minimum requirements for their students; and perhaps even volunteers who are working for others they do not know personally are inspired to give extra effort. Can this idea be formalized and empirically supported? Do people work harder when they are working for more than just themselves? In this paper, we attempt to find empirical evidence of this motivating power of others and pinpoint the conditions under which it can be an effective motivator.

BACKGROUND AND THEORY

Work Motivation

Research has shown that motivating workers is not as simple as paying them directly for their input or output (for reviews, see Prendergast 1999, Fehr and Falk 2002). Indeed, very few jobs only use pay-for-performance incentives and those that do usually require easily measurable output (such as assembly line workers or fruit pickers). Instead, firms rely on a mix of both extrinsic and intrinsic incentives to motivate their workers. Extrinsic incentives commonly used in organizations are monetary compensation and employee benefits (healthcare, housing, tuition reimbursement, etc.). Organizations also use many non-financial motives to motivate their workers. These may include non-financial but still extrinsic motivations such as the threat of punishment or termination, or more intrinsic motivations such as reliance on employees’ innate enjoyment of the work, desire to uphold workplace norms as part of an organization’s culture (i.e., Falk and Ichino 2006).
On the one hand, both lab and field studies have shown that pay-for-performance contracts can improve effort for simple, repetitive tasks such as installing automobile glass (Lazear 2000) or repeatedly pressing two keys on a keyboard (Ariely et al. 2009). On the other hand, a great deal of research has shown that contracts that rely on extrinsic motivation often fail to perform as classical economic theory predicts. For example, *explicit* contracts with direct disincentives for poor performance have been shown to have detrimental effects relative to *implicit* contracts—imcomplete contracts without specific incentives for performance—that leave workers a large degree of discretion over their effort (Fehr and Gächter 2000, Fehr and Schmidt 2000).

However, when both extrinsic and intrinsic motivations are present, as in most realistic situations, research has also shown that low levels of positive incentives can also decrease performance (e.g. Gneezy and Rustichini 2000, Heyman and Ariely 2004), an effect usually explained by the “crowding out” of intrinsic motivation (for review, see Deci et al. 1999). That is, when people derive innate satisfaction from the task, receiving pay for it crowds out this intrinsic motivation but does not replace it with sufficient extrinsic motivation due to the low levels of pay-for-performance. In a classic example, research on the overjustification effect suggested by self-perception theory found that when children who previously engaged in drawing for their own enjoyment were asked to draw for an extrinsic reward (a certificate), the amount of time they spent drawing afterwards was significantly decreased (Lepper et al. 1973).

In this paper, we argue that prosocial motivations (Rioux and Penner 2001, Grant et al. 2007, De Dreu and Nauta 2009) such as working overtime to help a coworker can also be a viable option to motivate workers.

**Prosocial Motivation**

More recent work has started to relax the assumption that employees are purely self-interested by incorporating possible prosocial motivations for expending effort (Batson 1987, Grant 2007). Prosocial motivations refer to those that encourage people to offer their own resources, whether time, effort, or money, for the benefit of someone or something other than themselves. Prosocial motivation has been differentiated from purely intrinsic motivation as effort driven by the desire to help others not based on “inherent interest in the work itself, but rather by introjected goals of avoiding guilt and protecting self-esteem or by identified goals of fulfilling core values and identities” (Grant 2008). That is, although prosocial motivation also does not lead to any material gain for the worker, it is nonetheless outcome-focused with a goal of benefiting others rather than the self (Grant 2007).

While there is building evidence suggesting that employees often work to benefit others and not only themselves (Thompson and Bunderson 2003, Meglino and Korsgaard 2004, Grant, Campbell, Chen, Cottone, Lapedis and Lee 2007, De Dreu and Nauta 2009), research has yet to compare how motivated
people are to work to benefit oneself versus work to benefit others. In order to do so, we first start off by reviewing what may theoretically drive prosocial behavior.

**Benefiting others as a goal of prosocial behavior.** As prosocial behavior is ultimately about helping others, a logical conclusion would be that it is motivated by underlying concern or regard of those they are helping. For example, perhaps people are altruistic, or actually care about others’ outcomes. Altruism is a feeling of wanting to help so that the prosocial behavior is fully the choice of the helper. Economists usually model altruism by adding others’ utility to the utility function to explain prosocial behavior as utility-maximizing (e.g., Becker 1974). Although altruism is a potential motivating factor in certain types of relationships, some research has argued that it is unlikely to play a major role in laboratory interactions with strangers (Levitt and List 2007).

Psychologists have also studied the highly related concept of empathy, defined as “a vicarious affective response to others” or “an affective response appropriate to someone else’s situation rather than one’s own” (Hoffman 1981). Empathy is a definite possibility as to why people help others: Batson (1987) refers to this notion as the “empathy-altruism hypothesis.” In support of this hypothesis, many researchers have found that feeling empathy when observing another in need can lead to increased helping behavior (Coke et al. 1978, Dovidio et al. 1990, Batson et al. 1995, Batson et al. 1995).

**Benefiting the self as a goal of prosocial behavior.** Although the observed outcome of prosocial behavior is helping others, it is also possible that some of its underlying motivations are actually more egocentric. A possible motivator of prosocial behavior may be impression management (Schlenker 1980). As prosocial behavior is the helping of others, there is automatically a component of reputation or others’ judgment. Whether the source of critique is the self or others, people prefer to be perceived as close to an ideal self as possible. Research has shown that people act to maintain a positive self-presentation (Bandura 1977, Reis 1981), even when they are the only observers of their behavior. That is, people are motivated by the desire to have a positive self-image, and even if actions are only personally observed, they may nonetheless add to their self-image. Positive self-presentation is a strong motivator (Baumeister 1982) and could be a reason for prosocial behaviors in situations where others’ opinions can be influenced.

In some situations, not engaging in prosocial behavior can have negative consequences, and these negative consequences can also be a motivator of prosocial behavior. Disappointing others can lead one to experience the negative emotion of *guilt* due to the belief that one has not met expectations that others may have had. The perceived source or sources of guilt could be the person one is responsible for or third parties and does not require actually seeing the lowered impression as long as a source is imaginable.

Another possible motivation for prosocial behavior is reciprocity. According to Fehr & Gächter (2000), reciprocity “means that in response to friendly actions, people are frequently much nicer and
much more cooperative than predicted by the self-interest model; conversely, in response to hostile actions they are frequently much more nasty and even brutal.” We focus only on positive reciprocity as a possible motivation for prosocial behavior. When one person helps another person, there is some chance that the beneficiary will return the favor in the future. This future gain can be the motivation behind helping another. Going one step further, indirect reciprocity, akin to the golden rule (Alexander 1987), is also possible. People would want another person working on their behalf to put forth his or her best effort, so they should likewise put forth their best effort for another person. Computational biologists have shown that cooperation purely due to indirect reciprocity is evolutionarily stable (Nowak and Sigmund 1998, Panchanathan and Boyd 2004). Dufwenberg and colleagues (2000) have also found suggestive evidence that indirect reciprocity might be stronger than direct reciprocity.

Self-interested motivations for prosocial behavior can also focus on its social implications. As members of society, people tend to pay attention to social expectations and norms of how one ought to behave (Elster 1989, Fehr and Gächter 2000). Adherence to these social standards could lead to the experience of positive emotion (e.g., pride) and deviation to negative emotions (e.g., shame or guilt) (Krupka and Weber 2008), as well as social acceptance or sanctions. Therefore, prosocial behavior might be the result of following a social standard rather than the desire to help another. Reciprocity could be viewed as a norm in addition to a potential gain for the worker as mentioned above. Researchers have proposed economic models to describe the role of norms such as fairness and reciprocity (Rabin 1993, Levine 1998, Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Charness and Rabin 2002, Falk and Fischbacher 2006).

One possibility is that people may feel obligation or duty to another person that results in behavior benefiting the other person in some way. Berkowitz and colleagues demonstrated evidence for this feeling of obligation by manipulating the dependency relationship between participants, thereby directly inducing feelings of responsibility (Berkowitz and Daniels 1963, Berkowitz et al. 1964). Participants performed a manual labor task while a confederate “supervised.” A participant’s performance either did or did not affect his supervisor’s payment, but never affected his own payment. Perhaps unsurprisingly, participants worked harder when their supervisors’ payment depended on their performance. The authors referred to this as “social responsibility.” This type of obligation might be present in a work setting where one group member feels obligated to help a fellow group member with a task, even though the helping member does not have any explicit responsibilities in the task.

“Prosocial Incentives”

Though all these possible motivations may exist as to why people may engage in prosocial behavior, intuition may suggest that, ceteris paribus, individuals will not exert as much effort when
working for others under purely “prosocial incentives” as they would when working for themselves. When using the term “prosocial incentives,” we refer to those benefits that one receives from choosing to act in a prosocial way, such as gains in one’s self-image; reciprocal gains. However, we propose that on the contrary, individuals—under the right conditions—will put in as much if not more effort when working on behalf of another as when working for oneself.

Although we examine this topic in an experimental setting, the question we aim to answer is one with wider applicability. As social beings, individuals constantly work and interact with others, and they must make decisions on how much or how little they are willing to do for another. Because their resources are limited, individuals must make sacrifices in order to benefit others. By exploring how much of their own resources people are willing to give to others, we hope to better understand the psychology behind prosocial behavior, and perhaps find ways to utilize these prosocial incentives to better motivate people.

The theoretical contributions of this paper are thus two-fold. First, it adds to the existing literature on worker motivation by exploring possible sources of additional motivation beyond personal incentives and giving insight into when prosocial incentives are motivating to individuals and to what degree. Second, this paper will add to the group work and productivity literature. The majority of people in today’s workforce do not work alone: we have teams, departments, groups, etc., and we are certainly affected by these situations. Prosocial incentives may be a way to better motivate groups in productivity and teamwork.

There are numerous examples from both real-world behaviors as well as from laboratory experiments of people’s willingness to exert effort on behalf of another person without direct benefits to themselves. Therefore, the prediction of some effort under prosocial incentives is hardly controversial. However, our question is how much effort. We use a novel prosocial incentives contract that relies purely on social motivations to motivate performance in order to compare pay-for-performance to prosocial incentives. Rather than have each worker’s performance directly affect his or her own pay, each worker’s performance under social incentives determines the pay of another person, or “beneficiary.” While the beneficiary could in practice be anybody, in our experiments, each participant’s output determines another participant’s pay. Thus, this contract uses the same structure as pay-for-performance incentives but shifts the target of the pay to a beneficiary.

**HYPOTHESES**

Under what conditions can prosocial incentives encourage workers to exert effort? Our experiments ask participants to work for the benefit of another (prosocial incentives) without receiving payment based on their performance. We expect that participants will try at least a bit when in this
situation, confirming that there are motivating factors besides direct pay-for-performance that can result in people exerting effort:

_Hypothesis 1: People will exert positive effort when motivated by purely prosocial incentives, despite not receiving direct pay-for-performance incentives._

If people can be motivated by factors other than direct incentives in prosocial situations, then we are interested in uncovering the conditions under which this is true. According to Self-Determination Theory (Ryan and Deci 2000), people require three needs for functioning and growth: needs for competence, relatedness, and autonomy. Grant (2008) builds on this theory, applying it to prosocial motivation, and states that “prosocial motivation is most likely to enhance persistence, performance, and productivity when it is accompanied by intrinsic motivation.” This intrinsic value could be due to the innate entertainment value of the task (Experiment 1) or due to the gains in self-image (Experiment 2). Because of crowding out, participants who are under direct incentives with intrinsic value will do less, especially since the payment is low. Based on these findings, we propose our second hypothesis:

_Hypothesis 2: When participants are asked to complete a task that has intrinsic value, those under prosocial incentives will do better than those under small levels of direct incentives._

We attempt to test the effect of social expectations (or simulated norms) by setting reference points for our participants. We base our fourth hypothesis on Heyman & Ariely (2004), who compared money-market relationships with social-market relationships. Money-market relationships are those in which monetary compensation is offered for effort, and social-market relationships are those in which no money is involved or there is a gift reward. They found that in money-market relationships effort was directly related to level of payment: low payment led to low effort, while high payment led to high effort. However, in social-market relationships effort seemed to be from altruistic motives and insensitive to payment amounts: the level of effort was the same regardless of payment amount. In our study (Experiment 2), participants working for their own pay are in money-market relationship, while participants working for another’s pay are in a social-market relationship. Heyman & Ariely (Heyman and Ariely 2004) state that mentioning money as payment for one’s efforts activates frames and norms related to a monetary-marketplace. We anticipate that participants whose effort is directly related to their payment will think in terms of the amount of money they will be making and therefore will pay more attention to the reference points, while participants whose effort helps others will not think of their effort in monetary terms. While Heyman & Ariely (2004) dealt with direct payment levels (none, low, or high), we apply their general idea to reference points. We tell our participants of a low, medium, or high reference point, and we anticipate that when participants are asked to exert effort for their own payment, they are responsive to reference points since their effort will directly relate to their payment and
reputation. However, when under social incentives, participants are insensitive to reference points and will exert a reasonable amount of effort.

Hypothesis 4: When working for direct incentives, participants are responsive to reference points, but when working in a social market, participants are not sensitive to magnitude.

Berkowitz & Daniels (1964) found that participants were more willing to help when they had been previously helped, and offers the reciprocity principle as a possible explanation for helping behavior. They posit that the participants felt some obligation to “pay back” the help they received. Furthermore, in their study, the participant receives help from a different party than they give help to, showing that this reciprocity principle does not have to be specific to a person, but rather can be generalized to a different person. In Experiment 4, we attempt to replicate this finding as well as compare reciprocity as a motivation with other possible motivations of prosocial behavior. We explore reciprocity as a potential motivator of prosocial behavior and predict the following:

Hypothesis 5: Reciprocity matters and will motivate workers to put in effort for another.

We also test the effect of the “social responsibility norm” by manipulating the degree of the social cost of violating the norm. We manipulated whether or not the participant will have to meet their beneficiary. Participants who expect to tell their partner their performance directly should feel that doing poorly on their partner’s behalf is more aversive than those who have the experimenter acting as a buffer. This buffer is similar in spirit to having an agent between you and the beneficiary (Hamman et al. 2010). We refer to these treatments as high versus low social cost. We predicted that other-pay participants would work harder with high social costs.

Our society looks positively on giving help to someone who is in need of help, a social norm, and if one decides to deny help to someone who is in need, this is in violation of an established social norm. The social cost is in the loss in social approval if the norm is violated (Cialdini 1991). We anticipate people will be more likely to give up the effort and resources to help another when they must actually meet the beneficiary in person, since the beneficiary’s potential disapproval is more salient, raising the social cost, than when they do not have to meet the beneficiary.

Hypothesis 6: Workers will work harder when they expect to meet their beneficiary due to a higher social cost of failure.

Another aspect of social cost due to violation of social expectations or norms is the degree to which a person can be identified for the act of following or violating a social norm. For example, as the number of people increases, people are less likely to help, a phenomenon called “diffusion of responsibility” (Latane and Dabbs 1975). The responsibility to act according to the social expectation of helping someone in need gets diffused among the group of people in the situation, and as the group gets larger, people feel less obligated to work. This supports the idea that when the likelihood of being
identified is high, people will tend to act according to the social expectation, but when the likelihood is low, people will not be motivated to exert effort.

*Hypothesis 7: When the social cost of failure is less, from not being identified, others’ knowledge of one’s performance does not motivate effort.*

**OVERVIEW OF THE PRESENT RESEARCH**

In a series of four experiments, we examined when prosocial incentives are motivating, especially conditions under which social incentives are more motivating than direct incentives and identify reasons why this may be the case. Experiment 1 directly compared prosocial incentives and direct incentives and also included treatments without any pay. Experiment 2 switched to a letter-typing task and explored the influence of performance standards by manipulating reference points. Experiment 3 increased the social cost of violating social expectations by having participants report their effort levels to the beneficiary in person. Finally, Experiment 4 attempted to tease apart the possible underlying motives for why people may work harder under prosocial incentives than when under direct incentives.

**Experiment 1: Anagram task**

Experiment 1 pitted “prosocial incentives” against the direct incentives of pay-for-performance. Prosocial incentives were implemented in a sequential design by having each participant’s performance determine the next participant’s payment. We used an anagram task similar to TextTwist in which participants were given four minutes to generate three to seven letter words using a subset of seven letters. With 140 possible words, we expected the number of words generated to be a reasonably linear function of effort.

In addition to the two paid treatments, we also included two unpaid treatments to better compare the relative contributions of the various components of motivation. One of the unpaid treatments tested whether differences between the paid treatments were because prosocial incentives participants work harder or because direct incentives participants work less hard. That is, this treatment tested for “crowding out” of intrinsic motivation, which would predict that self-pay participants will generate even fewer anagrams than even unpaid participants. A second unpaid treatment tested the contribution of reference points.

**Methods.** Two hundred internet participants completed the experiment. Instructions informed participants in the paid treatments that they would be paid in Amazon.com gift certificates. In order to receive payment, participants had to give their full names and email addresses. This request served to activate self-identity, thus amplifying the potential to feel shame and guilt. To keep this element constant,
the instructions for participants in the unpaid treatments also asked for this information but explained that “this information will be used only in case of follow-up.”

Participants in the paid treatments learned that their performance would either affect their own payment or “the next person to take the study’s” payment at a rate of 20 cents per anagram generated. Participants then learned the number of anagrams the previous participant had supposedly generated: “The previous participant, zach burns, generated 63 anagrams thus earning [you] $12.60.” This number was chosen to be fairly high in order to maximize the reference point for social incentives, but not so high as to seem implausible. The name was displayed in lower case to heighten realism because the experimenter presumably would not purposely program a lower case name. Participants may have also guessed that the next participant would similarly be informed of their name and performance and thus have stronger feelings of potential shame.

In the social incentives treatment, the previous performance determined the current participant’s pay. This manipulation set a salient performance reference point for what other participants might expect to receive. Generating less than that number of anagrams may be seen as a failure to satisfy social expectations. Note that deceiving participants was necessary in order to keep the information and norm constant across participants.

Instructions in the “unpaid with information” treatment also included the previous participant’s performance but without the earnings clause, whereas instructions in the “unpaid without information” treatment did not include any information about the previous participant. This latter treatment was designed to explore the effect of telling participants the previous player’s performance. We expected that removing this information would remove part of the self-presentation incentive due to giving one’s name because the next participant would not know the previous participant’s performance. It also removes the reference point and thus the high performance standard. We can thus separate the effect of this information from the effect of responsibility toward other participants.

After the payment instructions, all participants received detailed instructions for the anagram task. Participants were given four minutes to generate as many three to seven letter anagrams as possible from a set of letters, “a d e r s t w.” At the end of the four minutes, the experiment automatically advanced to a screen of demographic questions. A final screen debriefed participants as to the purpose of the study and, in all but the unpaid without information treatment, the nature and necessity of the “previous participant” deception.

**Results.** Participants on average generated 20 percent of the possible anagrams ($M = 26.5$ anagrams, $SD = 14.1$, $median = 24$, $min = 2$, $max = 87$). Most participants (189/200) submitted some invalid or repeated words, with a mean success rate of 78 percent ($SD = 14.3\%$, Median = 80%), suggesting that a better measure of effort level is the number of words participants submitted ($M = 34.0$, $SD = 13.2$, $median = 34$, $min = 2$, $max = 115$).
SD = 17.4, median = 31.5, min = 2, max = 111). Participants pressed an average of 154 keys while entering these words (SD = 84.4, median = 133.5), implying an average submitted word length between four and five letters. Table 1 shows the performance of participants in each treatment as measured by both submitted and valid anagrams. Because both measures were positively skewed with some outliers, we log transformed the data for subsequent analysis and geometric means as a summary statistic. We also provide the same results using non-parametric tests.

Table 1. Exp.1: Performance on anagram task.

<table>
<thead>
<tr>
<th></th>
<th>Submitted Anagrams</th>
<th></th>
<th>Valid Anagrams</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
<td>Median</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>Unpaid w/o info</td>
<td>27.8</td>
<td>14.1</td>
<td>27</td>
<td>24.2</td>
</tr>
<tr>
<td>Unpaid with info</td>
<td>37.6</td>
<td>19.7</td>
<td>32.5</td>
<td>32.0</td>
</tr>
<tr>
<td>Self-Pay</td>
<td>33.5</td>
<td>19</td>
<td>29</td>
<td>29.5</td>
</tr>
<tr>
<td>Other-Pay</td>
<td>37.4</td>
<td>14.4</td>
<td>35</td>
<td>34.6</td>
</tr>
<tr>
<td>Total</td>
<td>33.9</td>
<td>13.7</td>
<td>32.5</td>
<td>29.8</td>
</tr>
</tbody>
</table>

Relative to participants who earned money for themselves (a.k.a. self-pay), participants who earned money for the next participant (a.k.a. other-pay) submitted and generated five more valid anagrams (ts = 1.73 and 2.03, ps = .09 and .05, two-tailed t-tests on transformed values; zs = 1.82 and 1.95, ps = .07 and .05, Wilcoxon rank-sum test on untransformed values). If actual performances in the other-pay treatment had been used to calculate payment (rather than the 63 anagrams earning $12.60 for each participant), other-pay participants would have earned $1.08 (24.7%) more than self-pay participants.

Participants’ performance in the unpaid-with-info treatment was roughly midway between performances in the self-pay and other-pay treatments and not significantly different from either for submitted or valid anagrams (ts = .72 and .70 for submitted, ns; ts = .92 and .92 for valid, ns). Participants in the unpaid-without-info treatment performed worse than all other treatments on both submitted (ts = 1.86, 3.61 and 2.31, ps = .07, .001 and .02, for self-pay, other-pay and unpaid with info, respectively) and valid anagrams (ts = 1.32, 3.58 and 2.21, ps = .19, .001 and .03).

Discussion. In Experiment 1, participants whose performance determined the next participant’s payment performed better than participants whose performance determined their own pay. However, the fact that unpaid participants performed in between the two paid treatments suggests that the difference between social incentives and direct incentives is not a simple main effect. Rather, there seemed to be some crowding out of intrinsic motivation under direct incentives. Thus, two theoretically separate advantages of social incentives over direct incentives may have caused the performance difference in the paid treatments: 1) less crowding out of intrinsic motivation and 2) higher motivation due to impression
management. Although both effects are interesting, they can be disentangled by switching to a task that is not as intrinsically enjoyable. The anagram game was susceptible to crowding out because it was relatively entertaining and because fairly small amounts of money were at stake. Moreover, the unpaid-without-info treatment performed significantly worse than the unpaid-with-info treatment, suggesting that competition may provide additional intrinsic motivation. We will further explore this desire to beat the previous participant’s score by manipulating performance standards in Experiment 2.

Another potential problem with Experiment 1 is that the average pay was different across treatments. Whereas participants in the self-pay treatment had to earn their own payment, thus receiving a geometric mean payment of $5.44, all participants received $12.60 in the other-pay treatment. Thus, another explanation of the difference in effort levels may be that participants exerted more effort as a form of positive reciprocation for higher than expected wages or, in other words, gift exchange (Akerlof 1982). If participants treated the previous participant’s efforts as a fixed payment, the other-pay treatment could be interpreted as an implicit contract, which previous research has already demonstrated can be more motivating than explicit contracts (e.g. Fehr et al. 1997, Gneezy and List 2006). We can separate the motivating effect of implicit contracts by not informing participants of their pay before starting the task.

A final problem with Experiment 1 is that performance on the anagram task could be perceived to be indicative of intelligence or ability. A worker’s low output may reflect a lack of ability or laziness, both of which are personally and socially aversive. Thus, exerting high effort on another’s behalf is doubly important for impression management purposes because low output may communicate some combination of laziness, low ability, and selfishness. It is possible that participants in the other-pay treatment worked harder to avoid seeming incompetent to the next person rather than to avoid social disapproval. Although potential attributions for low output due to lack of ability or laziness are also possible for direct incentives, social incentives may make the possibility of such an attribution more salient because it provides a concrete source of evaluation: workers know their beneficiary will directly observe the product of their output. However, because we did not measure attention to others in any of our experiments, this mechanism remains a topic for future research.

**Experiment 2: Letter-typing with explicit reference points**

Although Experiment 1 demonstrated that Hypotheses 1 and 2 were true under certain conditions, it had various shortcomings. Experiment 2 addressed many of these shortcomings and also tested for the robustness of the effect by switching to a mindless letter-typing task. Importantly, this task was pretested to be highly effort-based and not very skill-based. Pretest participants did not find the task to be enjoyable and it thus avoids some of the issues arising from intrinsic motivation in the previous experiment. In addition, Experiment 2 precluded the gift-exchange alternative explanation by informing each participant
of the previous participant’s performance only after the latter participant completed the task. In doing so, the previous participant’s performance is no longer a salient fixed payment. Removing the potential for crowding out under direct incentives and gift-exchange under social incentives also obviates the unpaid treatments for this and all future experiments. Finally, deception about the previous participant’s performance was no longer required.

Experiment 2 also explored why social incentives were more motivating than direct incentives in Experiment 1. One key component in Experiment 1 was the usage of a salient performance reference point, the previous participant’s performance. Removing this reference point in the unpaid-without-information treatment made participants significantly less motivated, but Experiment 1 did not include comparison treatments to determine the effect of reference points for direct and social incentives. Therefore, Experiment 2 manipulated the reference point for the two types of incentives by providing participants the purported average performance from a previous study as a reference point. We expected that manipulating the reference point would affect effort under social incentives by providing a starting point for forming social expectations. Thus, we predicted that higher reference points would produce higher performance in the other-pay treatment, but that reference points would be irrelevant for this purpose in the self-pay treatment. However, reference points may also provide a competition motive for participants under direct incentives, so we remain agnostic about the effect of reference points on direct incentives.

Methods. Two hundred and fifty-nine internet participants learned that they had a chance to receive an Amazon.com gift card by email and entered their email address to begin. Participants then learned that their performance would either affect their own or “the next person to take the study’s” gift card size at a rate of 3 cents per letter-pair typed. The typing task required participants to alternate between pressing the “F” and “J” keys for two minutes. Extra keystrokes were recorded, but only alternating F’s and J’s counted toward payment and showed up on the screen. To facilitate tracking of progress, a counter at the bottom left of the screen kept track of how many letter-pairs the participant had already typed. Participants also learned they would find out their gift card size and the previous participant’s performance at the end of the experiment.

To manipulate the reference point, participants in the medium reference point treatment learned that “the average performance in the pretest was 444 letter-pairs.” In the low and high reference point treatments, the reference points were instead set at 344 and 529, respectively. These reference points corresponded to the actual mean, the 25th and the 75th percentile performances of a pilot study. On the task screen, the program reminded participants of the reference point in a box at the bottom right of the screen.

Results. Participants typed an average of 968 letters (SD = 243, median = 959, min = 2, max = 1664), correctly typing 459 letter-pairs (SD = 107, median = 465, min= 1, max = 774). Most participants
(93%) made at least some typing errors, with a mean error rate of 4.5 percent (SD = 4.8%, median = 3.0%, max = 27.2%). Participants’ error rates were positively correlated with the number of letter-pairs they typed (r = .23, p < .001) and did not vary by treatment or reference point. Error rates were fairly small and most errors seemed to be a result of trying harder without regard for accuracy. Because we were more interested in effort than in performance, we used total number of letters typed as a proxy for effort in all analyses.

Figure 1 shows the number of letters typed for each payment type and reference point. A 2 (pay: self vs. other) x 3 (reference point: low vs. medium vs. high) ANOVA showed that participants who earned money for themselves (self-pay) and participants who earned money for the next participant (other-pay) typed about the same amount of letters overall (Ms = 973 vs. 962, F(1,253) = .03, ns). However, performance varied significantly by reference point (Ms = 933 vs. 948 vs. 1016, F(2, 253) = 3.08, p < .05), and this difference was qualified by a significant interaction effect (F(2, 253) = 3.38, p < .05). Self-pay participants typed significantly more letters for higher reference points than for lower reference points (F(1, 253) = 10.34, p < .01) but other-pay participants typed the same number of letters regardless of reference point (F(1, 253) = .00, ns). Self-pay participants worked harder than other-pay participants at the high reference point (F(1, 253) = 5.15, p < .05), but insignificantly less hard for the low and medium reference points (F(1, 253) = .54 and 1.18, ns). The contrasts for the low reference point becomes significant (F(1, 244) = 4.25, p < .05) if age is added as a covariate in the ANOVA as a proxy for skill level (F(1, 244) = 79.33, p < .0001).

*Figure 1. Exp.2: Mean number of letters typed.*
Discussion. Although Experiment 2 did not replicate the main effect of social incentives being more motivating than direct incentives, differences between social incentives and direct incentives depended on the reference point provided. Self-pay participants were responsive to reference points, perhaps using them as a goal to beat, and increased their efforts for the higher reference point. More importantly, and contrary to Hypothesis 3, other-pay participants worked at the same level across reference points. This disregard for performance standards suggests that people under social incentives have a fixed cost of failing to live up to social expectations and thus work only hard enough to balance out this cost with the cost of additional effort. This failure to scale up with expectations is especially problematic because it suggests that social incentives may only be equally or more motivating than direct incentives when the pay rate is low. We will explore this issue in more detail in the general discussion.

Experiment 3: Consequences of failure to meet social expectations

Experiment 3 attempted to more directly manipulate the social costs of failing to meet social expectations. To do this, Experiment 3 moved from the internet to the laboratory, which allowed more social connectedness than over the internet and facilitates the use of a truly paired design. Experiment 3 also allowed us to test for robustness to experimental setting and participants and avoided potential problems such as selection bias and participant attrition.

The key manipulation was that participants were separated during the task but told they would be exchanging scores with their partner after the task either face-to-face or via the experimenter passing a note. Keeping in mind the results of the previous experiment, high social cost might also change the competitiveness of self-pay participants, but without any guiding theory, we did not make any a priori predictions.

Methods. One hundred and four walk-in participants were recruited individually at the University of Chicago’s Decision Research Lab for a 15-minute experiment that could pay up to $20. When a dyad was complete, an experimenter led them to a “neutral” room and introduced them to each other as “partners.” Following consent procedures, an experimenter read the instructions aloud to participants as they followed along on their own copies and then led them to two separate rooms on either side of the neutral room to complete the letter-typing task. In order to make incentives sharper than previous experiments, payment remained at zero for the first 400 letter-pairs and then increased three cents per letter-pair thereafter. This 400 mark also served to set a reference point. Although participants were unlikely to construe it as the average score in a pretest, as in Experiments 2 and 3, this number could be reasonably construed as the worst past performance.

After completing the task, each participant was asked to handwrite a note recording his or her performance on a slip of paper with the words “I typed _____ letter-pairs, thus earning [my partner]
$______,” depending on payment type. Participants also completed a short questionnaire that included a task in which a variety of possible reasons for exerting effort in the letter-typing task were listed and participants ranked between 3 to 10 reasons. At the end of this questionnaire, participants in the low social cost treatment gave their slips of paper to the experimenters to give to the other participant and were paid separately in their own rooms. In the high social cost treatment, the experimenter instead brought participants back to the neutral room, where they exchanged their slips of paper and received their pay.

Results. A total of 13 pairs in each treatment completed the experiment. One person did not understand the instructions for the typing task and was excluded from analysis. Participants typed on average 1158 letters ($SD = 192$, $median = 1162$, $min = 546$, $max = 1567$), thus correctly entering 542 letter-pairs ($SD = 81$, $median = 544$, $min = 272$, $max = 713$). Five participants failed to meet the minimum mark of 400 letter-pairs and instead received a consolation prize of one dollar. Counting these participants as zero pay, participants earned on average $5.67 ($SD = 2.94$, $median = 5.75$, $min non-zero = 1.00$, $max = 12.50$).

Figure 4. Exp.4: Mean number of letters typed.

Nearly every participant (98%) made at least some typing errors, with a mean error rate of 6.1 percent ($SD = 4.2%$, $median = 6.0%$, $max = 18.0%$). Participants’ error rates were positively correlated with the number of letters they typed ($r = .55$, $p < .001$) and slightly less so with the number of letter-pairs.
correctly typed ($r = .29, p < .01$). We used the total number of letters typed as a proxy for effort in all analyses.

Figure 4 shows the average number of letters typed in each treatment. A 2 (pay: self vs. other) x 2 (social cost: low vs. high) ANOVA revealed no main effects of payment type and social cost ($F(1, 99) = .0004$ and .03, ns) but a slight interaction ($F(1, 99) = 7.40, p < .01$). As predicted, other-pay participants worked significantly harder when they expected to meet their partner after the task than if they expected the experimenter to pass their performance information along ($F(1, 99) = 4.17, p < .05$). However, this higher social cost also led to marginally lower performance for self-pay participants ($F(1, 99) = 3.26, p = .07$). Other-pay participants outperformed self-pay participants when performance information was exchanged in person ($F(1, 99) = 3.54, p = .06$), whereas the opposite was true when performance information was passed by the experimenter ($F(1, 99) = 3.87, p = .05$).

As in Experiment 3, performance was recorded every five seconds. Table 3 displays the relative speeds in the two halves of the task. Across all treatments, there was improvement in the second minute (Rank-sum test $z = 9.60, p < .0001$). This improvement was bigger and more significant in the other-pay treatment. Linear regression of relative rate on interval show that speed increased even throughout the second minute ($β = .0064, t = 8.22, p < .001$) but at a faster rate for the other-pay treatment ($βs = .0042$ vs. .0087, $t = 2.87, p < .01$).

Table 3. Exp.4: Relative typing rates for certain relevant intervals.

<table>
<thead>
<tr>
<th></th>
<th>1st minute</th>
<th>2nd minute</th>
<th>$Δ_{(2^{nd} - 1^{st})}$</th>
<th>Before 400</th>
<th>After 400</th>
<th>$Δ_{(before - after)}$</th>
<th>$z$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Low</td>
<td>0.978</td>
<td>1.022</td>
<td>0.043</td>
<td>0.990</td>
<td>1.031</td>
<td>0.041</td>
<td>4.19</td>
<td>0</td>
</tr>
<tr>
<td>Self-High</td>
<td>0.997</td>
<td>1.003</td>
<td>0.006</td>
<td>1.001</td>
<td>0.997</td>
<td>-0.004</td>
<td>0.05</td>
<td>0.96</td>
</tr>
<tr>
<td>Other-Low</td>
<td>0.967</td>
<td>1.033</td>
<td>0.067</td>
<td>0.988</td>
<td>1.043</td>
<td>0.055</td>
<td>5.92</td>
<td>0</td>
</tr>
<tr>
<td>Other-High</td>
<td>0.970</td>
<td>1.030</td>
<td>0.061</td>
<td>0.987</td>
<td>1.040</td>
<td>0.053</td>
<td>4.77</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0.978</td>
<td>1.022</td>
<td>0.044</td>
<td>0.991</td>
<td>1.028</td>
<td>0.037</td>
<td>7.68</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: The z-scores are a Wilcoxon rank-sum test of the difference in relative typing rates before and after reaching the minimum payment point of 400 letter-pairs. Rates were calculated only for participants who reached it.

Table 3 also shows relatives rates for before and after reaching the 400 letter-pairs threshold for payment. To avoid biasing the “before” intervals, only participants who actually reached this point are included in this analysis. Self-pay participants under high social cost maintained the same speed after reaching the threshold, but participants in every other treatment significantly increased their speed after reaching the reference point. This increase may have been due to the highly salient climb of the payment display in the program.

Participants ranked on average 3.8 reasons for exerting effort on the task, with 72 percent ranking only three and 7 percent ranking all ten reasons. In order to accommodate participants ranking different
numbers of reasons, we only counted the percentage of participants who ranked each reason as their top reason or amongst their top three reasons. Self-pay participants ranked “wanting to earn as much money as possible” as their top reason for exerting effort on the task (75% 1st, 92% top 3), followed by “curiosity as to how well I could do” (12% 1st, 81% top 3) and “because I always give my best effort in anything I do” (6% 1st, 37% top 3). Other-pay participants ranked “aversion to disappointing the other person” as their top reason (33% 1st, 80% top 3), followed by “curiosity as to how well I could do” (22% 1st, 75% top 3) and “because I always give my best effort in anything I do” (16% 1st, 39% top 3). However, other-pay participants were not affected by high or low social cost for listing aversion to disappointing the other person as their top reason for effort ($M_s = 35\%$ vs. $32\%$, $\chi^2(1) = .005, ns$) or among their top three reasons ($M_s = 81\%$ vs. $80\%$, $\chi^2(1) = .04, ns$).

Discussion. Consistent with Hypothesis 5a, Experiment 4 found that requiring other-pay participants to exchange information about their performances face-to-face made social incentives more motivating. We hypothesized that this increased motivation was due to the increased social cost of failure to meet social expectations. This finding is particularly important for organizational purposes because workers in actual firms do not generally have the level of anonymity afforded to online participants and do have to confront each other frequently. The fact that the heightened sense of accountability increased motivation gives hope for application to actual workplaces. On the other hand, the anticipation of face-to-face confrontation also made direct incentives less motivating. This decrease could be due to participants not wanting to outperform their partners by too much or could be due to anxiety from having to directly compare their result with their partner’s. We again leave this unpredicted result as a topic for future research.

Experiment 5: Counting task

Experiments 1 through 4 presented various possible explanations for why people work harder for others in certain situations, including crowding out of intrinsic motivation by direct incentives; indirect reciprocity; higher motivation due to impression management when working for others; competitiveness when comparing own performance to others; and effects of social expectations on self vs. other pay. Some other explanations include altruism, guilt for negatively impacting another, and disappointment aversion. Experiment 5 attempted to disentangle these possible mechanisms by methodically manipulating the beneficiary of one’s efforts (self-pay, other-pay, fixed pay) and degree of social exposure (results not shared, results shared with next participant, results shared with third party). In addition, to increase the robustness of our results, we switched to a previously-validated number counting task (Abeler et al. 2011).
Design. Table 4 shows all seven conditions used in Experiment 5 in a partial three-way factorial design crossing the type of payment scheme, the beneficiary of the participant’s effort, and the degree of social exposure. These seven conditions allow us to isolate possible motivations for working harder.

Table 4. All seven conditions in Experiment 5. Empty conditions were not run.

<table>
<thead>
<tr>
<th>Who Effort Affects</th>
<th>How Pay is Determined</th>
<th>Who knows about your effort (and knows you worked for them)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Only self</td>
</tr>
<tr>
<td>Self</td>
<td>Own effort</td>
<td>1</td>
</tr>
<tr>
<td>Next Participant</td>
<td>Fixed</td>
<td>3</td>
</tr>
<tr>
<td>Nobody</td>
<td>Previous Participant’s Effort</td>
<td>5</td>
</tr>
</tbody>
</table>

Condition 7 is our control condition, where we are able to get a baseline reading on how much people are willing to do the task. Condition 1 allows us to see the effect of pay-for-performance compared to condition 7. Condition 2 was included to serve to get a baseline reading of the effect of having an audience for one’s own performance and/or competing with the previous participant’s results. The italicized condition in the top row is similar to condition 2, and therefore not necessary. Condition 3 tests for whether people are motivated by something independent of the presence of others to exert effort for another person, including altruism, self-impression management, or guilt aversion. Condition 4 looks at people’s effort with an audience to benefit another. Through condition 5 we can test whether indirect reciprocity is a key factor in motivating people to work harder for others. If people act based on the idea that since they are receiving the benefits of others’ efforts they should try hard as well, then participants in condition 5 will try harder than participants in condition 4. Condition 5 is the same as condition 4 except for the fact that the payment for participants in condition 5 is dependent on the previous participant. We included condition 6 in our design to determine whether a difference exists between the effect of the person affected (the next participant) knowing and the effect of a non-affected other person knowing. The two italicized conditions in the bottom row get at the effect of an audience on performance, but this was captured by condition 2, and therefore these two conditions were left out of our design.

The counting task we used for this experiment consisted of counting the number of 1’s in randomly generated 15 x 10 grid of 0’s and 1’s. We chose a new task from the previous experiments because we found participants still found the letter-typing task to be enjoyable. By using the counting task, we were aiming to eliminate all intrinsic motivation from enjoying the task itself. Participants were shown one grid at a time, and after each grid, the participant was given the option to either continue onto the next grid or end the task. We did not place a time limit on how long a participant could spend on each grid since by doing so, we would have introduced another possible goal of accuracy within a time limit to the decision to try on the task.
Methods. One hundred and thirty-seven participants from the Columbia Center for Decision Sciences’ Virtual-Lab database completed the experiment. Participants began the task by clicking the link in their customized e-mail invitations. Participants were restricted to completing the experiment only if they received an invitation, completed the experiment once, and completed it in good faith.

After completing an electronic consent form, participants first filled out the Positive Affect Negative Affect Scale (PANAS) in order for us to measure their current affective state. They were then given general instructions on the task, explaining that they would be presented with grids containing 0’s and 1’s, and their assignment was to count the number of 1’s in each grid. The participants were told that accuracy counted, and that answers within one of the correct count would be considered correct and add 100 points to their score. After these initial instructions, participants were given two practice trials. These practice trials gave the participant a chance to see what the task was before reading the next set of directions and making a decision on how hard to try on the task. They also served as baseline for accuracy and speed in completing the task, since they appeared before payment instructions.

After the practice trials, the participants saw a second set of directions regarding payment depending on condition as shown in Table 4. Participants found out that their payment would be determined by either 1) his or her own performance, 2) by the performance of the previous person in the study, or 3) that it was pre-determined. It also indicated what their performance would affect either 1) their own payment, 2) the next person’s payment, or 3) that it only affected the number of points they received but was unrelated to payment. Participants for whom performance was consequential were told that their points would be converted to payment at the rate of $.25 per 100 points (for each correct answer). Finally, participants were also told what feedback they would receive at the end of the task. Participants in the “only self” conditions in Table 4 were simply told they would see their results and earnings after finishing the grids. Participants in “next participant” conditions were told they would see their own as well as the previous person’s results. Finally, participants in “experimenter” condition would see only their own results, but to induce the idea that another person would see their results other than the person benefiting from their effort, they were also told that a research assistant would review their performance. Importantly, all participants found out the amount of their payment (and of the next participant’s payment in the “next participant” conditions) only after completing the task in order to avoid having participants use the amount of their payment (or the next participant’s payment) as anchors to base their effort on. After the payment instructions, participants were asked to estimate how much they thought they would make in the task, and if applicable, how much they thought the next person would make.

Participants then finally began the paid trials of the counting task. They could spend as long as they wanted to count each grid and, after submitting their answer for each grid, were given the choice to continue to the next grid or to stop counting and see their results. They could continuing counting grids as
long as they wanted to continue but were told that there was an upper limit. However, in order to avoid anchoring their effort levels, we did not tell participants the limit of 32 grids. After choosing to stop or reaching the limit, participants filled out the PANAS again and then were shown a results page showing each grid they completed, whether the counted correctly for each grid, how many they answered correctly in total, and the amount of payment they would receive. Depending on the conditions as mentioned before, some participants also saw the same set of results from the previous participant in addition to their own. Participants finished the study by answering questions regarding their motivations and thoughts on the task, as well as demographics and a manipulation check.

**Results.** Participants whose efforts determined their own pay and whose results were not seen by others (condition 1) worked significantly harder than participants who were paid a fixed amount and whose results were not seen by others (condition 7) ($F(1,168) = 5.40, p = .02$). Participants whose efforts determined their own pay, saw the previous person’s results, and the next participant saw their results (condition 2) did not differ in performance from participants in condition 1 ($F(1, 168) = 0.00, p = .95$). Participants who received a fixed payment but whose effort determined the next person’s payment without anyone else knowing (condition 3) did not differ in performance from participants who received a fixed payment while working for the next person’s payment with the beneficiary viewing their results (condition 4) ($F(1,168) = 0.00, p = .99$). Participants in condition 3 did not differ in effort from condition 7 ($F(1,168) = 0.30, p = .58$). Participants worked somewhat harder when their payment was determined by the previous participant, their effort was paying the next participant, and the beneficiary saw the participant’s results compared to participants whose payment was fixed, their effort was paying the next participant, and the beneficiary saw the participant’s results ($F(1,168) = 3.34, p = .07$).

**General discussion**

In this paper, we have shown that we can motivate people to work on behalf of someone else without any direct benefit to themselves. We have argued that this motivation derives from the desire to meet social expectations and maintain positive social and self-image. In a series of four experiments, we have shown that social incentives can be used to motivate workers, that they are sometimes more motivating than direct incentives, especially when the output benefits concrete beneficiaries who the worker expects to meet.

Experiment 1 found evidence that social incentives can be more motivating than direct pay-for-performance incentives but found that unpaid participants also generated more anagrams than those under direct incentives, suggesting that the relatively small pay direct incentives crowded out intrinsic motivation. We suspect that part of this crowding out was due to choosing a task many participants considered fun for competition. Social incentives may have also benefitted from “gift exchange” due to
the fact that these participants essentially received a fairly high fixed payment upfront. An additional treatment removed previous performance information and led to significantly fewer anagrams, thus providing additional support for the idea that competition made the task fun. This no information treatment also raised the possibility that reference points could affect motivation.

Experiment 2 more directly manipulated performance standards by providing participants with one of three values as the “average performance from a pretest.” This manipulation was intended to show that participants under social incentives would use this reference point as a starting point for what the next participant might expect from them. Experiment 2 was also able to remove the “gift exchange” motivation by not providing information about the previous participant’s performance before the task and reduced concerns about crowding out by using a less intrinsically motivating letter-typing task. Contrary to our initial hypothesis, we found that performance under social incentives did not vary across the different reference points. On the other hand, participants under direct incentives seemed to set a goal to beat the reference point, thus working harder for higher reference points. Social incentives were more motivating than direct incentives for the low reference point but less motivating for the high reference point.

We designed Experiment 3 to clarify the mechanism for social incentives by varying the determinedness of the beneficiary. We predicted that participants paired with the previous participant would find their beneficiary more tangible and thus feel worse for failing to meet social expectations than if paired with the undetermined next participant. We found the predicted stronger effect of social incentives on behalf of more determined beneficiaries but also found that the opposite was true for direct incentives: participants paired with the previous participant stopped soon after passing the reference point provided and thus performed worse than those paired with the next participant. This unpredicted effect may have been due to heightened feelings of competition with an undetermined participant relative to a determined one, similar to the lessened excitement of watching a recorded sports competition without knowing the results.

Although Experiment 3 was consistent with our proposed mechanism, we designed Experiment 4 to more directly manipulate the social cost of failing to meet expectations. We ran this experiment in the laboratory with a truly paired design and manipulated social costs by having participants in the high social cost treatment report their score to their partner in a face-to-face meeting. With social incentives, this meeting allowed participants to potentially experience their partner’s disappointment with a low payment, and we predicted that the anticipation of potential shame would increase participants’ efforts relative to those who only had to pass a note via the experimenter. As predicted, we found that social incentives were more motivating when social costs of failing to meet expectations were higher. However, we again found an unpredicted effect in the opposite direction for direct incentives. Participants who expected to
report their earnings in person seemed reluctant to perform too well, as evidenced by the lack of a boost in typing rate upon reaching the minimum payment threshold found in the other treatments.

**Do social incentives scale up?**

We proposed that workers are willing to work on behalf of another person out of a desire to avoid looking lazy or selfish, whether to themselves or others. That is, we posited that social incentives are motivating because people try to avoid paying the purely social cost of failing to meet social expectations. Although we were able to make this cost more salient by making the beneficiary more concrete and could increase the cost with a face-to-face meeting, Experiment 2 suggests that it is more of a fixed cost that may not scale up with the level of expectations. Although other research suggests that we could further increase the social cost by making behavior more observable (e.g., Levitt and List 2007) or self-identity more salient (Wicklund 1975), this fixed cost will nonetheless always be balanced against the variable cost of additional effort. For example, it is highly unlikely that a participant in one of our experiments would be willing to run a marathon so that another participant could be paid a few more dollars (cf. Olivola 2010). Therefore, it is possible that our experimental results are due to having found a “sweet spot” at which the cost of tarnishing of one’s social and self-image is higher than the cost of exerting effort in the task. That is, social incentives may not be more motivating than direct incentives for tasks that require higher costs of effort.

Another consideration is that although we did not focus on the role of altruism in our experiments, feeling good for helping others may nonetheless have contributed to the effectiveness of social incentives. Without manipulations of anonymity or observability, it is difficult to separate the effects of altruism from those of impression management (Levitt and List 2007). However, it is still useful to consider whether altruism can scale up with higher costs of effort. If we treat altruism as literally having the beneficiary’s outcome as part of the worker’s utility function, then it should motivate workers under social incentives in essentially the same way pay-for-performance directly incentivizes workers. That is, the more the worker produces under social incentives, the more pay the beneficiary gets, which in turn makes the worker happier, just like receiving more pay under direct incentives. This type of mechanism would scale up with higher pay rates and higher costs of effort, but does not seem particularly realistic. Instead, the most likely obstacle to altruism is more steeply diminishing marginal returns to helping others earn more, relative to earning more for oneself. Therefore, we expect altruism, as a mechanism for social incentives, to have the same problems to scaling up with higher costs of effort.

However, social incentives’ inability to fully scale up with higher costs of effort does not necessarily make it worse for those harder tasks, since direct incentives also rely on the relationship between marginal utility versus marginal cost. It remains an empirical question as to how social and direct
incentives compare for more effortful tasks outside our experimental setting. For example, would such an arrangement work with real apple pickers or factory workers over the course of eight hours?

**Considerations for use in organizations**

Although scaling up to higher effort tasks is one pressing question, there are many other issues to consider if we are to actually institute social incentives in an organizational setting. Although the experiments we have presented all required real effort and actually paid for performance, we readily admit that generalizing from experimental effort to worker motivation in organizations is a major leap. In this section, we review some of the potential challenges to applying social incentives in an organizational setting. Due to the limitations of our online and laboratory experiments, we discuss these issues as areas for future research.

**Repetition.** Perhaps the most important question for practical organizational use is whether social incentives can remain motivating over longer periods of time. In our experiments, participants performed a novel task under a novel incentive scheme for a single, short period of time. It is therefore possible that novelty may have contributed to the effectiveness of social incentives in the experiments and that this novelty could wear off after initial exposure. People in real world jobs could quickly tire of working on behalf of others. They may also tire of letting the fate of their own earnings lay in the hands of others. People tend to be risk averse and may require that social incentives be paid at a higher rate than equivalent direct incentives in order to compensate for the uncertainty in pay. We can test whether social incentives can remain motivating over time in a laboratory setting simply by having multiple rounds of work. However, even this would be difficult to generalize from since experimental participants generally do not rely on their earnings as a main source of income. A field experiment using temporary workers working for real pay may be more helpful for this purpose (e.g., Gneezy and List 2006).

**Job complexity.** Another issue, although one that is less theoretically interesting, is that most jobs, including essentially all “white collar” ones, are by nature multifaceted. Whereas apple pickers have the single measurable objective of picking apples, it is more difficult to measure performance in more complex jobs. When a worker’s performance cannot be quantified, it is impossible to set explicit pay-for-performance contracts. In addition, paying for performance on a single dimension can create perverse incentives to maximize that dimension at the detriment of other job functions (Prendergast 1999). It is similarly challenging to apply social incentives to complex jobs where output is hard to measure, but because this is a problem with any form of non-fixed incentive, we point out this limitation but offer no ways to circumvent it.

**Group production.** Relating to the previous point, many jobs also require people to work together in groups. Even if output is quantifiable, it may be difficult to isolate any given individual’s
unique contribution to the group’s performance. Again, this inability to measure performance is a limitation for any non-fixed incentive. Direct incentives can at least pay group members on the basis of the group’s total performance, though at the risk of social loafing (Darley and Latané 1968, Karau and Williams 1993). However, social incentives are unlikely to be effective on the group level, for example if one group’s performance determines another group’s pay. Not only are groups less concrete than individuals (Slovic 2007), the risk of social loafing may be even more problematic, as there is even less accountability to another group’s members than one’s group members.

On the other hand, it is worth noting that this paper can also be viewed as a contribution to the literature on group incentives (for review, see Shepperd 1993) by providing a “micro” foundation for why very small groups can be effective. That is, even in the minimal group of two people where only one worker has any incentive to exert effort, that worker can nonetheless be motivated to work under social incentives.

**Relationship between effort and performance.** Another obstacle to using social incentives in organizations also relates to tying payment to worker performance. Whereas the experiments in this paper tried to use tasks with a direct mapping from effort to performance, relaxing this condition may weaken the effectiveness of social incentives. For example, suppose that luck is a major factor in the production function, such that high effort does not always lead to high performance. If social incentives are mostly driven by impression management motives, then the possibility of a bad outcome despite high effort could be highly aversive. Without any other way to signal high effort, workers may choose to simply not work at all. Of course, this aversion could be true for direct incentives as well.

Another related variable is the extent to which a worker’s skill matters for performance. In jobs where skill varies to a great degree, relatively high-skilled workers who are aware of their skill may be less likely to work hard to give their beneficiary high payment when their own benefactor is likely to have lower skill than them and thus is less likely to provide them with as high a payment. On the other hand, relatively low-skilled workers may likewise be unwilling to work hard when they know their performance will be disappointing even with high effort. This issue is particularly troublesome if we consider the selection effects of allowing workers to choose social incentives versus direct incentives. High-skilled workers may should rationally chose direct incentives since it should provide them with more pay on average than social incentives when paired with a low-skilled worker. Low-skilled workers conversely may chose social incentives expecting to increase their pay by pairing with a higher-skilled worker but find their plan backfiring due to no high-skilled workers choosing social incentives.

**Resistance to implementation.** These selection effects are part of a broader issue with resistance to implementation of social incentives due to managers’ and employees’ fear of the unknown. A compromise that may be more palatable to employees and managers alike is a hybrid contract with both
direct and social incentives. This type of contract maintains the familiar appeal of relating harder work to more money but also allows social incentives to further increase motivation levels beyond what individual workers deem satisfactory. This may also have the additional benefit of giving people who subscribe to the norm of self-interest an excuse to work on behalf of others (Miller 1999). That is, social incentives may not appeal to workers who have the opposite impression-management goal: to appear self-interested.

Another obstacle to implementation is that even pay-for-performance contracts are used in a fairly limited range of jobs such as farming, manufacturing, and sales. However, even some salary-based contracts include both fixed and performance-based components with the latter usually in the form of bonus payments. Therefore, we could implement social incentives for just the performance-based component of these contracts. However, Epley and Gneezy (2007) argued that a bonus may be perceived as a windfall gain, which suggests that workers may have low expectations for bonus relative to regular pay. This difference in perception could pose a problem for social incentives since if bonus payments are considered extra, people may believe that others also have low expectations for bonuses and therefore feel less compelled to meet work hard on their behalf. That is, their lowered social expectations would lead to lower effort.

Organizational benefits

Although there are some serious challenges to overcome to applying social incentives in an organizational setting, the starting point of this paper is that social incentives can be more incentivizing than direct incentives. Assuming that we can satisfy the conditions required for this to be true, social incentives can also benefit organizations in ways beyond higher motivation. For example, if workers know which of their coworker’s performance determines their pay, they may be more willing to help that coworker or monitor his or her effort. Although both of these behaviors may lead to inefficiency if done to excess, they are generally desirable for firms. Motivating coworkers to monitor each other’s efforts serves a particularly useful function, one that is usually quite costly for organizations to formally implement, particularly for specialized jobs in which measuring effort is difficult for those without specific knowledge (e.g., information technology firms).

Social incentives can also lead to perceptions of a friendly firm culture, whether due to the fact that everyone is working on behalf of someone else or due to actual coworker helping. This firm culture can in turn improve worker satisfaction (Odom et al. 1990), which has many downstream consequences such as higher organizational commitment and employee retention (Currivan 1999). In addition, workers may be motivated to like the coworker responsible for their pay, which would improve workplace relationships.
CONCLUSION

Classical economic theories model people as perfectly selfish and lazy. No right-minded *homo economicus* (Thaler 2000) should ever give away money or do any unmonitored work when not directly receiving pay for the output. Yet people *do* give to anonymous strangers in experiments and charities in real life, and they *do* exert effort without any marginal benefit in salary-based jobs. They are even willing to exert effort on behalf of others without any direct reward for themself. This paper has shown that when people are made responsible for another’s welfare, these social incentives can be highly motivating. As a result of wanting to avoid violating social expectations, *homo sapiens* appear munificent, working hard on behalf of others, and, under the right conditions, even harder than they work for themselves.
References


