PAY ENOUGH OR DON'T PAY AT ALL*

URI GNEEZY AND ALDO RUSTICHINI

Economists usually assume that monetary incentives improve performance, and psychologists claim that the opposite may happen. We present and discuss a set of experiments designed to test these contrasting claims.

We found that the effect of monetary compensation on performance was not monotonic. In the treatments in which money was offered, a larger amount yielded a higher performance. However, offering money did not always produce an improvement: subjects who were offered monetary incentives performed more poorly than those who were offered no compensation. Several possible interpretations of the results are discussed.

I. INTRODUCTION

According to standard economic reasoning, an increase in the financial incentives provided for an activity will improve performance. This prediction is a conclusion of very basic assumptions in economic theory: performance is positively related to effort; effort is unpleasant, and money is good. We should therefore observe a monotonic and increasing relationship between monetary compensation for an activity and the performance level of that activity.[1] The main aim of this paper is to provide a test of this prediction, in a controlled laboratory environment, which pays particular attention to the comparison between the total absence and the presence of monetary rewards. Our main result is that performance varies in a nonmonotonic way with incentives.

The monotonic relationship predicted by the theory may fail in concrete situations, either in real life or in experiments, because factors different from money and effort may enter into the decision of the agent. For instance, a person may be reluctant to work for very small compensation because this fact might signal his general willingness to accept a small wage, and thus weaken his future bargaining position. A different reason, which is more commonly suggested in the economics literature, is that people follow social norms that prescribe a behavior independently of any

* We thank Colin Camerer, Gary Charness, Douglas Dejong, Martin Dufwenberg, Ernst Fehr, the editor Edward Glaeser, Georg Kirchsteiger, Matthew Rabin, Robert Rosenthal, Ariel Rubinstein, two anonymous referees, and an anonymous editor for helpful comments and suggestions. Special thanks go to Eric van Damme for his comments as well as for his help in raising the money needed for the experiments.

1. This argument requires that changes in compensation are small enough so that the income effects are negligible as compared with the substitution effect. In our experiments and in the literature we discuss, this is always the case.

© 2000 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

The Quarterly Journal of Economics, August 2000

monetary compensation. Donating blood may be considered a duty to the community that one should perform when possible. A monetary compensation may destroy this sense of duty and produce a net decrease in the action.[2] A different social norm that may be undermined by monetary compensation is reciprocity. Suppose that an action is originally performed in return for a previous benefit, but that money is paid for it. Then the compensation rather than the reciprocity will probably be taken as a motivation for that action. The incentive for reciprocity is destroyed, and the action becomes less appealing on its own merits.[3] As we will see, none of these explanations seems adequate for our results.

The issue of the effect of rewards on behavior has been debated in psychology throughout the past four decades. Behaviorist theory had the same opinion as standard economics, although for completely different reasons. According to instrumental conditioning, reward offered for an activity which is in itself neutral or even mildly unpleasant, will eventually associate a positive valence to that activity. So in the long run a past reward has a positive effect on the performance of that activity.[4]

This conclusion of behaviorist psychology was challenged at the beginning of the seventies by the cognitive psychology school. They offered an alternative view: an activity has a motivation of its own, independent of any reward, called *intrinsic motivation.* A reward, different from this intrinsic motivation (in particular, but not only, a monetary reward) may replace the intrinsic motivation.[5] The net effect may be a reduction of the overall motivation,

2. Titmuss {1970} claimed that monetary compensation might undermine the sense of civic duty. He considers the specific example of blood donation in Titmuss {1971}, where he argues that the introduction of monetary compensation will make the *quality* of blood donated worse. Arrow {1972} discusses his thesis: he predicts that an increase in price will eventually produce an increase in supply. More recently, the work of Frey and several coauthors {Frey 1994; Frey, Oberholzer-Gee, and Eichenberger 1996; Frey and Oberholzer-Gee 1997} has presented and defended the idea that price incentives may crowd out motivation. Kohn {1993a, 1993b} has criticized incentive plans because they make people less enthusiastic about their work.

3. In the experiments of Fehr, Gachter, and Kirchsteiger {1996} the introduction of explicit incentives reduced the performance of workers in a firm-worker relationship because the norm of reciprocity was compromised. This point is discussed in Fehr and Gachter {1998}, and more extensively, with additional evidence, in Fehr and Rockenbach {2000}. In his field study on management behavior, Bewley {1995, 1997} notes that real-life managers know well that it is not wise to depend on financial incentives alone as motivators.

4. A clear exposition of this point of view is in Skinner {1953}.

5. Definitions and measurement of intrinsic motivation are still controversial: but a basic condition for the existence (and empirical evidence) of intrinsic motivation is that the activity should be exercised even when reward is absent.

and hence a reduction of the activity itself. We can formulate the same idea in the language, more familiar to economists, of preferences: if the reward directly affects the utility of an individual in a negative way (because it reduces the intrinsic motivation), then performance may decline with the increase in monetary incentive.[6] The main conclusions of these studies were that positive rewards, in particular monetary rewards, have a negative effect on intrinsic motivation. If a person is rewarded for performing an interesting activity, his intrinsic motivation decreases. The negative effect is significant only if the reward is contingent on the performance; subjects who are paid a fixed positive amount, independent of their performance, do not display reduction in intrinsic motivation.[7]

1. Two Experimental Tests

In this paper we test experimentally the effects of monetary incentives on performance. In the first experiment a group of 160 students at the University of Haifa were asked to answer a set of 50 questions taken from an IQ test. The students were paid a fixed amount of 60 NIS (New Israeli Shekel[8]) for participating in the experiment. They were divided into four different groups, corresponding to four different treatments. The first group was simply asked to answer as many questions as they could. To subjects in the second group we promised an additional payment of 10 cents of a NIS per each question that they answered correctly. To subjects in the third group we promised an amount of 1 NIS, and to those in the fourth group an amount of 3 NIS per question. We observed that the average number of questions answered correctly declined from slightly more than 28 in the first group to 23 questions in the second. The number increased to more than 34 in the third group, and was stable at 34 in the fourth group. As we

The thesis was suggested in Deci {1971}, and further discussed among others in Deci {1975}, Deci, Cascio, and Krusell {1973}, and Kruglansky, Alon, and Lewis {1972}.[5] A rather large set of experiments showed that a lowering indeed occurred: an early overview of this literature and its experimental evidence is in Lepper and Greene {1978}.

6. This is the model of motivation crowding-out, presented in Frey and Oberholzer-Gee {1997}, for instance. Their model is discussed in Section IV below.

7. A consensus opinion in experimental psychology is far from being reached. Cameron and Pierce {1994} and Eisenberger and Cameron {1996} have provided meta-studies on the topic of the effect of rewards on motivation, evaluating more than two decades of studies on the issue. They also provide a useful review of the literature we have discussed. The final conclusion is still unclear: for instance, they find a negative effect of tangible rewards, and a positive effect of verbal rewards.

8. At the time of the experiment, NIS 3.5 = $1.

argue in the next section, where the details of the results are presented, the effect was statistically significant.

Our second experiment involved high school children, again in Israel, who were doing volunteer work. Every year, on a predetermined day, students go from house to house collecting monetary donations that households make to societies for cancer research, assistance to disabled children, and so on. In our experiment, we divided 180 of these students into three groups. The first served as a control group: we simply gave a small speech to the subjects recalling the importance of the activity they were going to perform. To the second group, in addition to the speech, we made a promise to pay 1 percent of the total amount collected. To the third we promised to pay 10 percent of the amount collected. In both cases it was made clear that the payment was financed by us, and not by the societies. In this second test we observed that the amount collected was smaller in the second than in the first group. The average amount for the third group was higher than in the second group, but still lower than in the first. Again, the results were statistically significant.

Since monetary incentives, at low values, would appear to have a detrimental effect on performance, we investigated whether the subjects were aware of this. In a final set of experiments we asked subjects to decide what incentive they would provide to other subjects working on their behalf. In this experiment subjects in the first group were paid according to the performance of the subjects to whom they were giving the incentive. They could choose between a no-reward and a low-reward. The incentive they decided to pay was subtracted from their payoff. The majority chose the low incentive. This incentive was more costly, and as we have seen was inducing a worse performance, so it is the wrong contract in the principal-agent relationship.

Our experiments focus on performance, a matter of central interest for economics. This variable also provides an objective, quantitative measure of the effect. In particular, we study the differential effect of small and large rewards. In contrast, the experimental psychology literature, motivated by the distinction between intrinsic and extrinsic motivation, has different measures of the intrinsic motivation to be the dependent variable. The measures commonly used are the amount of time freely spent on the activity and the report by the subjects on the motivation.[9]

There is a second difference between our study and the

9. These are the two variables that Cameron and Pierce {1994} and Eisenberger and Cameron {1996} consider in their meta-studies.

experiments performed in the psychology literature. Psychologists study behavior *modification* through conditioning (in the case of the behaviorist school) or learning (for the cognitive school). We do not. To illustrate the difference, we may consider the classic experiment reported in Deci {1971}. He had college students play with a puzzle in three successive sessions. In the first session participants were left to play freely. In the second session subjects in one group received payment if they solved the puzzle, while the control group did not. In a third session the subjects were again left to play freely. The amount of time spent on free activity in the first and third session was taken as a measure of intrinsic motivation. Deci found that in the third session the experimental group spent less time than the control group playing with the puzzle, and he concluded that the reward offered had decreased the intrinsic motivation of subjects in the first group over the three sessions. We study the behavioral response to different rewards in a single-stage setup. The comparison is across individuals, not across successive periods for the same individual following the reward.

Finally, we test both the effects of the introduction of reward, and the effects of an increase in the reward. In experimental economics the question is "how much closer to the prediction of economics and game theory does an increase in monetary rewards bring the behavior of subjects?" The focus is on changes in rewards, always keeping these positive.[10] We broadened this to allow for a start-up effect and showed that the overall pattern can be nonmonotonic, thus accounting for experimental evidence in psychology that seemed inconsistent with the findings of the experimental economics literature.

II. The Experiments

1. The IQ Experiment: Design

The experiment was conducted at the University of Haifa. The subjects were 160 male and female undergraduate students

10. Frey and coauthors have conducted field studies of the effect on intrinsic motivation from the point of view of economists. In particular, Frey and Oberholzer-Gee {1997} study answers to a questionnaire asking people about their attitude toward a nuclear waste repository in their region. Their study is based on stated behavior in hypothetical choice situations, in an interview face to face. Since nobody likes to be considered greedy, the inconsequential statements might be biased by the desire to show off. The questions were asked sequentially of the same subjects. The dependent variable was not performance, but willingness to exchange the inconveniences of nuclear waste against money. In their review, Smith and Walker {1993} find that increasing rewards brings the behavior closer to the predictions of economic theory, and reduces the variance around the mean. In the more recent review, Camerer and Hogarth {1999} conclude that the evidence is more controversial.

from all fields of study, with an average age of 23. The subjects took part in the experiment divided into four different groups of 40 students each, corresponding to four different treatments that we describe below.

At the beginning of the experiment, each student was promised a fixed payment of NIS 60 for participation. They were then told that the experiment would take 45 minutes, and they would be asked to answer a quiz consisting of 50 problems taken out of a psychometric test used to scan applicants to the university. This test is similar to the GMAT exam: the participants were told that this was a sort of IQ test. The problems in the quiz were chosen to make the probability of a correct answer depend mostly on effort. In particular, emphasis was placed on questions involving reasoning and computation rather than general knowledge.

In the four different treatments subjects were promised different additional payments for each correct answer. In the first group no mention was made of any additional payment. In the second group subjects were promised an additional payment of 10 cents of a NIS per question answered correctly. The amount promised was of step 1 NIS and 3 NIS, respectively, for the third and fourth group.

After the introduction, the quiz was distributed. Participants were not allowed any material on their tables except the quiz itself, and were told that only those who stayed until the end of the experiment would be paid. No clarifying questions by students were allowed during this time. At the end of the experiment participants were told where and when to go to collect their earnings. The instructions are presented in Appendix 1.

2. The IQ Experiment: Results

Appendix 2 reports the number of correct answers for each subject. Summary statistics are presented in Table I.

The average number of correct answers was 28.4 out of 50 questions in the first group. The average was 23.07 in the second group, where subjects were getting an additional 10 cents per correct answer. The average was then higher, even compared with the first group, and equal to 34.7 in the third group (one NIS), and 34.1 in the fourth group (three NIS).

A nonparametric Mann-Whitney U test based on ranks can be used to investigate whether the sample of correct answers came from populations with the same distribution. In Table II we report the results of a pairwise comparison of the different treatments. A

TABLE I

SUMMARY STATISTICS FOR THE IQ EXPERIMENT, FOR THE DIFFERENT TREATMENTS
The Lower Fraction is the Fraction of Subjects Who Gave a Number of Correct Answers Less than 16

	No payment	10 cents	NIS 1	NIS 3
Average	28.4	23.07	34.7	34.1
Standard deviation	13.92	14.72	8.88	9.42
Median	31	26	37	37
Average top 20	39	34.9	42.35	41.6
Standard dev. top 20	5.25	6.79	3.63	4.18
Average bottom 20	17.8	11.25	27.05	26.6
Standard dev. top 20	11.56	10.22	5.07	6.82
20th quantile	40	35	44	43
80th quantile	20	0	26	25
Lower fraction	15%	27.5%	0%	0%

TABLE II

MANN-WHITNEY *U* TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS OF MEDIANS OF CORRECT ANSWERS BY TREATMENT

	No payment	10 cents	NIS 1
10 cents	.0875	—	—
NIS 1	.0687	.0004	—
NIS 3	.0708	.0006	.6964*

(Prob. $> |z|$, where z is the test statistic). An asterisk indicates that for that comparison we cannot reject (at a .9 level of significance) the hypothesis that the two samples come from the same distribution.

number in the intersection of row and a column indicates, for the corresponding pair of treatments, the probability of getting at least as extreme absolute values of the test statistic as we observe, given that the two samples come from the same distribution.

The difference between the distributions in the zero payment and in the low payment is significant, at a .9 level of significance. The difference between the distributions in the high payoff treatments (1 NIS and 3 NIS) is not significant. Finally, the distributions in these latter treatments are significantly higher than the distributions in the case of the zero and 10 cents marginal payoffs. For instance, the *p*-values for the comparison between the 1 NIS treatment and the zero and 10 cents payment are .0687 and .0004, respectively.

The difference among treatments persists if we compare subgroups with similar performance. For instance, we ranked the

TABLE IIIa
MANN-WHITNEY *U* TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS OF MEDIANS OF CORRECT ANSWERS BY TREATMENT, USING THE BEST 20 OBSERVATIONS OF EACH TREATMENT

	No payment	10 cents	NIS 1
10 cents	.0381	—	—
NIS 1	.0146	.0003	—
NIS 3	.0339	.0008	.4222*

TABLE IIIb
MANN-WHITNEY *U* TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS OF MEDIANS OF CORRECT ANSWERS BY TREATMENT, USING THE 20 WORST OBSERVATIONS OF EACH TREATMENT

	No payment	10 cents	NIS 1
10 cents	.0299	—	—
NIS 1	.0120	.0000	—
NIS 3	.0270	.0001	.7552*

top half and the bottom half of the participants according to performance, and we then tested the significance of the difference between distributions. The results are presented in Tables IIIa and IIIb.

A similar comparison can be made between the distribution of the top ten participants. The difference between the 1 NIS and 3 NIS is clearly not significant. The difference between the 10 cents and 1 NIS as well as 3 NIS is significant.

The diversity in individual performance may be due to differences in several different factors, such as skill, general knowledge, and preferences for money and effort. Our results seem to indicate that the effect of the introduction of monetary incentives and their change affect in the same way individuals with different characteristics, as higher talent, or higher willingness to put out effort.

4. The Donation Experiment: Design

To illustrate the next study a premise is necessary. In Israel a few "donation days" take place every year. Each of these days is devoted to a society that collects donations from the public for some purpose, such as cancer research, disabled children, etc. High-school students go from door to door to collect the donations.

Normally, the students are organized into groups according to the class in which they study, and each group is then divided into pairs of students who work together as a team. Each pair receives a certain number of coupons, which serve as receipts for the donors. The amount collected by each pair on the donation day depends mostly on the effort invested: the more houses they visit, the more money they collect. This is especially true because the students do not have to "sell" the donation, since most people are already familiar with it from television announcements and advertisements.

We had 180 high-school students around the age of 16 participating with three treatment levels. There were two groups of participants, each with fifteen couples, for each treatment. Each pair received coupons amounting to NIS 500 altogether. In the discussion that follows we report jointly the results for the two groups at each treatment level.

In the first treatment an experimenter appeared before each of the groups and told them about the importance of the donation they were to collect, and that the society wished to motivate them to collect as much money as possible. They were told that the results of the collection would be published, so that the amount collected by each pair would become public knowledge. The second treatment was conducted similarly: but after the same speech, each pair was promised 1 percent of the amount that the two of them collected. Finally, in the third treatment each pair was promised 10 percent of the amount they collected.

In the second and third treatments it was made clear that the payment was made from funds additional to the donation, provided by us, and that the societies would receive the total amount of the donation as usual. The activity of collecting donations then went on as usual, according to the procedure described above.

5. The Donation Experiment: Results

The precise amount collected by each group for the three different treatments is reported in Appendix 3. We report the most important summary statistics in Table IV. The average amount collected was 238.67 over 500 for groups in the first treatment (with no payment). The average fell to 153.67 in the second group. It was 219.33, higher than in the second treatment (but lower than the first) in the third treatment.

To test the significance of these results, we use the nonparametric Mann-Whitney U test based on ranks to investigate

TABLE IV
SUMMARY STATISTICS FOR THE DONATION EXPERIMENT, FOR THE DIFFERENT TREATMENTS

	No payment	1 percent	10 percent
Average	238.6	153.6	219.3
Standard deviation	165.77	143.15	158.09
Median	200	150	180
Average top 20	375.33	272	348
Standard deviation top 20	111.92	98.64	110.46
Average bottom 20	102	35.33	90.66
Standard deviation bottom 20	66.13	52.08	63.97
20th quantile	100	0	50
80th quantile	450	250	400

TABLE V
MANN-WHITNEY *U* TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS OF MEDIANS OF AMOUNTS OF MONEY COLLECTED BY TREATMENT

	No payment	1 percent
1 percent	.0977	—
10 percent	.7054*	.0515

whether the amounts of money donated came from the same distribution. The results of the test are reported in Table V.

The difference between the average collection in the first and in the second group is significant, at a .9 level of significance. When the payoff increased to 10 percent of the amount collected, the average collection was 219.33. The amounts collected in this treatment were significantly higher than the amounts collected in the 1 percent treatment, but not significantly higher than the amounts collected when no payoff was given.

We compared the top ten collections in each treatment. The difference between the amounts collected in the 1 percent treatment and the amounts collected in the other two treatments was significant. As in the IQ experiment, this result indicates that the difference between treatments is uniform among subjects with high and those with low performance.

III. HOW SUBJECTS PERCEIVE THE EFFECTS OF MONETARY INCENTIVES

The evidence we have presented seems to indicate that the effect of monetary incentives can be, for small amounts, detrimen-

tal to performance. Since incentives are usually designed to affect performance in an optimal way, it is interesting to investigate whether this effect is anticipated. To address this question, we conducted a test based on both the IQ and donation experiments. In both cases subjects were promised a payment proportional to the performance of a different person (their "agent"), and they had to choose the incentive scheme for the agent.

1. The IQ Experiment with Principal and Agents

The subjects in the experiment were 53 students in the role of principals. They were told that they would be matched with another player. They were given a short introduction in which they received an explanation of the task that their "agents" would perform, namely, answering questions from the admission test that the subjects took in the IQ experiment. They were then told that each one of them would be paid 1 NIS for every correct answer given by his agent.

The principals had to choose whether a payment of 10 cents of NIS or a zero payment was to be made to their agent for every correct answer. This payment would be paid out of the amount of 1 NIS paid to the principal. The principals were told that the agent would know in advance how much he was going to be paid for every correct answer, but that he would not know that the principal had to decide first whether to pay him nothing or 10 cents. He would not even know that a principal existed. This was the only decision the principals had to make. At the end of the experiment participants were told where and when to go to collect their earnings.

Out of the 53 subjects 46 subjects, a proportion of 87 percent, chose to pay 10 cents for every correct answer of the agent. With this choice they reduced their income in two ways: by providing a payment to the agent, and by reducing the performance of the agent because of the negative effect of low rewards.

2. The Donation Experiment with Principal and Agents

In the donation experiment we also had a group of students who played the role of "principals." They were told that they would be matched randomly with one pair who had already collected money, and would be paid 5 percent of what this pair had collected. The principals had to decide whether they wanted us to choose the pair from the group that did not receive any payoff or

from the group that received 1 percent of what they have collected. The payment to the agent was made out of the 5 percent.

The results confirmed what we observed in the previous test. Out of the 25 participants, 19 (that is, a proportion of 76 percent) preferred to be matched with an agent who was paid 1 percent of the amount he collected.

IV. Interpretation of the Results

1. Summary of the Results

The following facts seem to be clear from the analysis of the results. If we compare the treatment in which monetary compensation was not even mentioned with the one in which it was, then we may conclude that the monetary compensation produces a reduction in the performance. But in the set of treatments in which a monetary compensation is offered, a higher monetary incentive produced a higher performance. This result indicates a discontinuity at the zero payment of the effect of monetary incentive: for all positive but small enough compensations, there is a reduction in performance as compared with the zero compensation, or, better, with the lack of any mention of compensation. Also, subjects in the same population as those who exhibit this behavior do not have a clear or strong perception of this discontinuity. They seem to indiscriminately apply the rule, which is valid in the region of positive compensations, that higher payment provides higher performance.

2. Intrinsic and Extrinsic Motivation

A first interpretation of these results is possible along the lines provided by cognitive psychologists: the activity has an intrinsic motivation, and the introduction of a monetary reward, which is an extrinsic motivation, displaces the first, and the net effect may be a reduction in the activity. We have to see whether this explanation is justified in the case of our experiments.

A first version of the interpretation is that the subject is weighting intrinsic and extrinsic motivation, and the intrinsic motivation is reduced directly by the monetary compensation. A model presenting this version formally is provided in Frey and Oberholzer-Gee {1997}. In this model the agent has a utility function of the activity a and the monetary reward r. The function has one term, $u(a,r)$, with the standard features, added to a second

term describing the intrinsic motivation, $m(a,r)$. This latter term captures the negative effect of monetary reward on intrinsic motivation, because the derivative of m with respect to r is negative. The optimal level of activity as a function of the reward r is then computed in the usual way. Comparative static analysis shows that the optimal activity may be decreasing in r.

This model is insightful, and has the virtue of reducing the discussion to a special case of standard economic analysis. However, it seems unable to explain the discontinuity: a small change from zero to a very low payment should reduce (if the utility function has curvature) a small change in the activity.

A second version is based on a cognitive explanation of the displacement, based on attribution theory along the lines first suggested in Festinger {1957} and Heider {1958}. This explanation of the displacement is presented in Bem {1965, 1967} and Kelley {1967, 1971}. For instance, Bem suggests that people interpret their actions as any outsider does, by looking at the reasons and motives for what they do. If a subject observes himself doing an action when no exogenous compensation is provided, then he interprets his motive as an intrinsic motivation. If a monetary compensation is provided, then the same subject will interpret his motive as being the monetary reward.

In this version of the interpretation the discontinuity is explained: the introduction of small compensation has the effect of changing the perception, and this change is large, independent of the amount. But this explanation is only appropriate in sequential experiments, where a subject has the opportunity of observing himself acting for some incentive, modify his self-perception, and then change his behavior. It is less appropriate in our single-stage setup. The next interpretation seems to us the most convincing.

3. Incomplete Contracts

The contract describing the experiment is an incomplete contract. All successive instructions given by the experimenter provide partial completion of its terms. When no monetary compensation for the correct answers is provided, the contract is probably interpreted as "A payment of 60 NIS is provided for participation in the experiment, and I now know that this participation consists of answering the questions." Some (but not all) of the subjects felt that it was their side of the bargain to answer those questions.

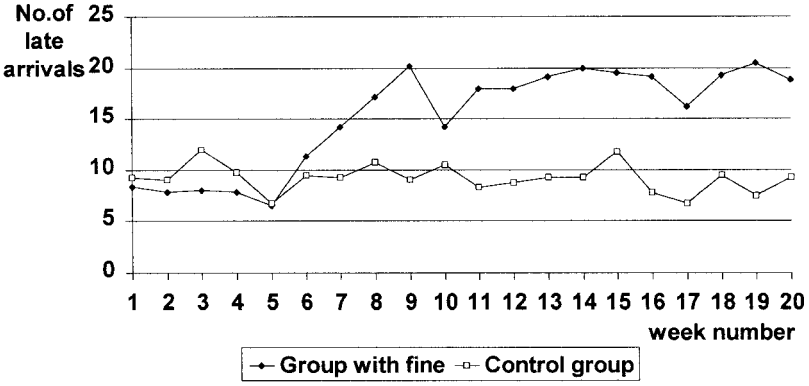

FIGURE I
Average Number of Late-Arriving Parents Each Week, by Group Type

The introduction of additional compensation for correct answers changes the perception of the contract: "Sixty NIS were paid for showing up. The activity of answering the questions is now paid by rate." This rate is now the reference point to decide the appropriate intensity of the activity. When the rate is low (10 cents), the activity is low: there is no necessary connection with the level of activity provided in that different situation, where answering questions is "due."

An important prediction given by our explanation (but also by the second version of the previous interpretation) is that the change in perception, once realized, is hard to reverse. This is indeed the case in a related field study {Gneezy and Rustichini 2000} that we conducted on the effect of penalties. The study was conducted in a group of day-care centers, where parents were coming later than the due time to collect their children. In the test group we introduced at the fourth week of the study a fine for late arrival. The fine was 10 NIS for a delay of ten minutes or more. The fine was cancelled at the seventeenth week. The effect was an increase in the number of late arrivals after the introduction of the fine. In Figure I we report the average delays for the control and the test group, in the twenty-week period.

This result is consistent with a shift in perception of an incomplete contract, as in the interpretation we have just provided. Further support for this explanation comes from the behavior of parents after the seventeenth week: the number of

late arrivals stayed constant, as one would expect after the perception of the contract has changed.

An interesting extension of the IQ experiment designed to test this theory would be to take the participants of the IQ quiz experiment, where 10 cents has been offered per each question answered correctly, and in a second experiment ask them to perform the same task, for no compensation. We conjecture that the performance will be worse than the one of subjects who have been offered zero compensation from the start, and of course worse than the performance of the same participants in the first experiment.

4. How Small Is a Small Amount?

In our experiments the subjects who were paid 10 cents for 1 NIS for each right answer gave a worse performance than did those who were not given any payment. Ten cents may sound like very small compensation, almost insulting and therefore the practical implications of our findings minimal; but two qualifications are necessary. First, not all small compensations may be considered insulting. For example, consider the practice of paying back a small amount for recycling a soft drink bottle, which is common in many European countries. Anecdotal evidence suggests that people are less willing to recycle when this small compensation is offered, than they are in places where no money is offered, and failure to recycle is simply considered bad behavior. It is unlikely that the amount offered is considered insulting. A different explanation in this case might be that people are afraid of looking "cheap" for making the effort of recycling to collect the small amount.

Second, insulting compensations are not necessarily small compensations. The reason is that the amount of money offered changes the perception that people have of "what the contract is about." In particular, it may not be safe to assume that adding an incentive leaves the utility of the other incentives unchanged. A certain amount of monetary compensation may be perceived as too small when compared with the other relevant factors, even if it is not too small in itself. We may think of real-life situations where a nontrivial amount of money may sound disproportionately small compared with other factors. For instance, an increase in salary by $200 per month to a professor, as compensation for a smaller

office, may be worse than no compensation. Similarly, small honoraria for seminar speakers may be counterproductive.[11] This factor is likely to be more important when factors like health or reputation are at stake. So while in our experiments it is clear that "too small" is somewhere between 10 cents and 1 NIS, the exact determination of this quantity in experimental or real-life situations is likely to be difficult and subtle.

Our theory has an interesting implication for experimental economics. It seems widely accepted now that a sufficiently high reward is sufficient to reduce the variance of the observed behavior around the mean value of the behavior as predicted by the theory. The seemingly natural implication that a small reward simply produces higher variance might be false: at the low end of the scale of rewards, there might be paradoxical behavior, of the type observed in our experiments. In fact, the behavior with small payment may be, as it is in our experiments, more distant from the prediction than the behavior with zero proportional payment. This may be important in cases where large payments are impossible, for example, because of ethical or legal reasons. The rule that "a small payment is better than nothing" might be a bad rule.[12]

5. The Two Experiments

There is of course an important difference between the two experiments. In the donation study there is an intrinsic motivation that is clearly identifiable: the altruistic reason that is, after all, the motivation for the students before they became subjects in our experiment. This difference, for instance, may explain a significant difference in behavior between the two studies. In the IQ test the performance with a substantial payment (1 NIS) reaches a level well above the one in the treatment with zero payment (34.7 correct answers against 28.4). In the donation study, even a payment of 25 NIS per person does not bring the performance back to the level achieved in the zero payment group (219.3 NIS against 238.67). Of course an even larger payment might be enough. So the issue of a systematic difference in behavior in the two environments seems interesting, and the point is worth further study.

11. We thank an anonymous editor for suggesting this point.
12. A similar point is argued in the review by Camerer and Hogarth (1999).

V. Conclusions

In this paper we have provided quantitatively precise evidence, in a controlled environment, of the effect of the introduction of monetary compensation on performance, which includes a precise comparison of the cases in which the reward was given in different quantities or not given at all. The result has been that the usual prediction of higher performance with higher compensation, *when one is offered,* has been confirmed: but the performance may be lower *because of* the introduction of the compensation.

On the basis of this precise evidence we may begin the search for a satisfactory explanation. Further research, in theory and experiments, is necessary, and we have indicated some of the promising directions. In the meantime, the most convincing explanation seems to us to be based on cognitive arguments: contracts, social or private, are usually incomplete, and regulate an interaction in a situation of incomplete information. The introduction of a reward modifies some of the terms of the contract, but also provides information. The new behavior produced by the contract is a response to the combination of a new payoff structure and the new information. The difficulty is that the standard Bayesian updating of information seems unsuited for this situation.

Appendix 1: Instructions for the IQ Experiment

Introduction

The instructions are simple, and if you follow them carefully you may earn a considerable amount of money. The experiment will take about 45 minutes.

In the experiment you are asked to answer a quiz of 50 problems taken from a psychometric test used to scan applicants to the university. It is a sort of IQ test.

You will be paid NIS 60 for showing up to the experiment. [*The following sentence was not included in treatment 1:* "In addition, you will be paid NIS 0.1" (in treatment 2, NIS 1 in treatment 3, NIS 3 in treatment 4) for every correct answer you give.]

The money will be paid to you, privately and in cash, at the end of the experiment.

Do you have any questions?

APPENDIX 2: THE NUMBER OF CORRECT ANSWERS GIVEN IN THE IQ EXPERIMENT BY PARTICIPANTS ACCORDING TO TREATMENTS

Obs. #	No payment	Obs. #	10 cents	Obs #	NIS 1	Obs. #	NIS 3
1	**49**	41	**50**	81	**49**	121	**50**
2	**48**	42	**44**	82	**47**	122	**50**
3	**48**	43	**44**	83	**47**	123	**47**
4	**45**	44	**43**	84	**46**	124	**45**
5	**42**	45	**40**	85	**46**	125	**44**
6	**42**	46	**39**	86	**45**	126	**44**
7	**42**	47	**36**	87	**44**	127	**44**
8	**40**	48	**35**	88	**44**	128	**43**
9	**37**	49	**35**	89	**44**	129	**42**
10	**37**	50	**35**	90	**43**	130	**41**
11	**37**	51	**34**	91	**41**	131	**41**
12	**37**	52	**34**	92	**41**	132	**39**
13	**36**	53	**32**	93	**41**	133	**39**
14	**36**	54	**32**	94	**40**	134	**39**
15	**36**	55	**31**	95	**40**	135	**38**
16	**35**	56	**30**	96	**38**	136	**38**
17	**34**	57	**26**	97	**38**	137	**37**
18	**34**	58	**26**	98	**38**	138	**37**
19	**34**	59	**26**	99	**38**	139	**37**
20	**31**	60	**26**	100	**37**	140	**37**
21	**31**	61	**24**	101	**34**	141	**37**
22	**31**	62	**23**	102	**33**	142	**36**
23	**31**	63	**23**	103	**33**	143	**36**
24	**29**	64	**22**	104	**33**	144	**34**
25	**29**	65	**21**	105	**31**	145	**33**
26	**24**	66	**21**	106	**31**	146	**31**
27	**23**	67	**21**	107	**30**	147	**31**
28	**23**	68	**19**	108	**29**	148	**31**
29	**23**	69	**19**	109	**29**	149	**28**
30	**22**	70	**13**	110	**29**	150	**27**
31	**22**	71	**11**	111	**28**	151	**26**
32	**20**	72	**8**	112	**28**	152	**25**
33	**20**	73	**0**	113	**26**	153	**25**
34	**18**	74	**0**	114	**23**	154	**21**
35	**7**	75	**0**	115	**22**	155	**20**
36	**3**	76	**0**	116	**22**	156	**20**
37	**0**	77	**0**	117	**22**	157	**19**
38	**0**	78	**0**	118	**21**	158	**19**
39	**0**	79	**0**	119	**20**	159	**17**
40	**0**	80	**0**	120	**17**	160	**16**

APPENDIX 3: THE AMOUNT OF MONEY COLLECTED BY STUDENTS IN THE DONATION EXPERIMENT ACCORDING TO TREATMENTS

Obs. #	No payment	Obs. #	1%	Obs. #	10%
1	**0**	31	**0**	61	**0**
2	**0**	32	**0**	62	**0**
3	**0**	33	**0**	63	**20**
4	**40**	34	**0**	64	**20**
5	**80**	35	**0**	65	**30**
6	**100**	36	**0**	66	**50**
7	**100**	37	**0**	67	**100**
8	**120**	38	**0**	68	**100**
9	**120**	39	**0**	69	**120**
10	**130**	40	**30**	70	**140**
11	**150**	41	**50**	71	**150**
12	**150**	42	**80**	72	**150**
13	**150**	43	**100**	73	**150**
14	**190**	44	**120**	74	**150**
15	**200**	45	**150**	75	**180**
16	**200**	46	**150**	76	**200**
17	**240**	47	**150**	77	**200**
18	**250**	48	**180**	78	**240**
19	**250**	49	**210**	79	**250**
20	**300**	50	**230**	80	**250**
21	**330**	51	**240**	81	**290**
22	**340**	52	**240**	82	**290**
23	**350**	53	**250**	83	**350**
24	**420**	54	**250**	84	**380**
25	**450**	55	**250**	85	**400**
26	**500**	56	**300**	86	**410**
27	**500**	57	**330**	87	**460**
28	**500**	58	**400**	88	**500**
29	**500**	59	**400**	89	**500**
30	**500**	60	**500**	90	**500**

TECHNION, HAIFA, ISRAEL
BOSTON UNIVERSITY

REFERENCES

Arrow, K. J., "Gifts and Exchanges," *Philosophy and Public Affairs,* I (1972), 343–362.

Bem, D. J., "An Experimental Analysis of Self-Persuasion," *Journal of Experimental and Social Psychology,* I (1965), 199–218.

——, "Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena," *Psychological Review,* LXXIV (1967), 183–200.

Bewley, T., "A Depressed Labor Market as Explained by Participants," *American Economic Review, Papers and Proceedings,* LXXXV (1995), 250–254.

——, "A Depressed Labor Market as Explained by Participants," manuscript, Department of Economics, Yale University, 1997.

Camerer, C. F., and R. M. Hogarth, "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," 1999, California Institute of Technology Discussion Paper No. 1059, in *Journal of Risk and Uncertainty,* XIX (1999), 7–42.
Cameron, J., and W. D. Pierce, "Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis," *Review of Educational Research,* LXIV (1994), 363–423.
Deci, E., "Effects of Externally Mediated Rewards on Intrinsic Motivation," *Journal of Personality and Social Psychology,* XVIII (1971), 105–115.
____, *Intrinsic Motivation* (New York and London: Plenum Press, 1975).
Deci, E., E. L. Cascio, and J. Krusell, "Sex Differences, Verbal Reinforcement and Intrinsic Motivation," Paper presented at the Meeting of the Eastern Psychological Association, Washington, May 1973.
Eisenberger, R., and J. Cameron, "Detrimental Effects of Reward. Reality or Myth?" *American Psychologist,* LI (1996), 1153–1166.
Fehr, E., S. Gachter, and G. Kirchsteiger, "Reciprocity as a Contract Enforcement Device," *Econometrica,* LXV (1996), 833–860.
Fehr, E., and S. Gachter, "Reciprocity and Economics: The Economic Implications of *Homo Reciprocans,*" *European Economic Review,* XLII (1998), 845–859.
Fehr, E., and B. Rockenbach, "The Hidden Costs of Economic Incentives. Contractual Contingencies Crowd Out Voluntary Cooperation and Reciprocity," University of Zurich Discussion Paper, 2000.
Festinger, L., *A Theory of Cognitive Dissonance* (Evanston, IL: Petersen, 1957).
Frey, B. S., "How Intrinsic Motivation Is Crowded in and out," *Rationality and Society,* VI (1994), 334–352.
Frey, B. S., F. Oberholzer-Gee, and R. Eichenberger, "The Old Lady Visits your Backyard: A Tale of Morals and Markets," *Journal of Political Economy,* CIV (1996), 1297–1313.
Frey, B. S., and F. Oberholzer-Gee, "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-out," *American Economic Review,* LXXXVII (1997), 746–755.
Gneezy, U., and A. Rustichini, "A Fine Is a Price," *Journal of Legal Studies,* XXIX (2000), 1–18.
Heider, F., *The Psychology of Interpersonal Relations* (New York City, NY: Wiley, 1958).
Kelley, H. H., "Attribution Theory in Social Psychology," *Nebraska Symposium on Motivation,* XV (1967), 192–238.
____, *Attribution in Social Interaction* (New York City, NY: General Learning Press, Module Series, 1971).
Kohn, A., *Punished by Rewards* (Boston, MA: Houghton Mifflin, 1993a).
____, Why Incentive Plans Cannot Work, *Harvard Business Review,* LXXI (1993b), 54–63.
Kruglansky, A. W., S. Alon, and T. Lewis, "Retrospective Misattribution and Task Enjoyment," *Journal of Experimental Social Psychology,* VIII (1972), 493–501.
Lepper, M. R., and D. Greene, *The Hidden Costs of Reward: New Perspectives in the Psychology of Human Motivation* (Hillsdale, NJ: Lawrence Erlbaum Associates; New York: 1978).
Skinner, B. F., *Science and Human Behavior* (New York, NY: Macmillan, 1953).
Smith, V. L., and J. M. Walker, "Money Rewards and Decision Cost in Experimental Economics," *Economic Inquiry,* XV (1993), 245–261.
Titmuss, R. M., *The Gift Relationship* (London, UK: Allen and Unwin, 1970).
____, "The Gift of Blood," *Trans-action,* VIII (1971); now in *The Philosophy of Welfare,* Selected Writings by R. M. Titmuss, B. Abel-Smith and K. Titmuss eds. (London, UK: Allen and Unwin, 1987).