Relative Thinking and Task Performance: Does a Larger Fixed Payment Reduce the Perceived Magnitude of the Pay-For-Performance Component?

Ofer H. Azar*

Ben-Gurion University of the Negev

Keywords: Relative thinking, Experimental economics, Compensation schemes, Task performance, Behavioral economics, Judgment and decision making, Pay-for-performance.

JEL codes: M52, C91, J30, D01, L29.

*Ofer H. Azar, Department of Business Administration, School of Management, Ben-Gurion University of the Negev, P.O.B. 653, Beer Sheva 84105, Israel. Tel.: +972-8-6472675; Fax: +972-8-6477691. E-mail address: azar@som.bgu.ac.il. I thank Colin Camerer, James Dana, Stefano Della Vigna, Ernst Fehr, Daniel Friedman, Uri Gneezy, David Laibson, David Levine, George Loewenstein, Ulrike Malmendier, Robert Porter, William Rogerson, and participants in seminars given at the University of Copenhagen, Interdisciplinary Center Herzliya, the Israel Economic Association 22nd annual conference and the 2006 ESA European Regional Meeting for helpful conversations and comments. I am grateful to Galit Dori, Tamar Kugler, and Michal Barry for valuable research assistance. Financial support from the Russell Sage Foundation and the Phillipe Monaster Center For Economic Research at Ben-Gurion University of the Negev is gratefully acknowledged.

Abstract

Many previous experiments suggest that people exhibit "relative thinking": they often consider relative price differences even when only absolute price differences should matter. The article reports the results of an experiment that tests whether relative thinking exists in a context never explored before, of task performance with mixed compensation schemes that include both fixed and pay-for-performance components. Such compensation schemes are prevalent in many occupations (for example salespeople and managers) and therefore the article addresses an important practical issue. Surprisingly, relative thinking disappears in this context: the ratio between the pay-for-performance compensation and the fixed compensation does not affect effort. To test whether the different context or the introduction of financial incentives (which were not used in previous studies of relative thinking) is the reason that relative thinking disappears, a hypothetical condition where subjects make similar decisions but without incentives was run. Relative thinking was not documented, suggesting that in the context of task performance people do not exhibit relative thinking regardless of financial incentives. The article therefore contributes both to the literature on relative thinking and to the area of personnel economics and designing incentive schemes in firms.

1. Introduction

Many experiments conducted over the last three decades suggest that people often consider relative price differences even when only absolute price differences should matter, a behavior that was recently denoted "relative thinking" (Azar, 2004).[1] The seminal experiment in this literature is the one reported in Tversky and Kahneman (1981). They asked people whether they were willing to drive 20 minutes in order to save $5 on a calculator when they were going to buy a calculator and a jacket. When the calculator's price was $15 and the jacket's price was $125, 68 percent of the subjects were willing to drive, but when the calculator's price was $125 and the

[1] Much of the previous literature refers to this behavior as "mental accounting." Mental accounting refers to a situation where people do not treat all their money as one resource, but rather split it into different mental accounts (see Thaler, 1985, 1999). There are several reasons, however, why "relative thinking" seems a better terminology to use. First, "mental accounting" is used to describe many other behaviors in addition to relative thinking, for example when someone is happy to receive a gift from his spouse that he did not want to purchase himself, although they have a joint bank account (Thaler, 1985), or when someone is unwilling to purchase another theater ticket when he loses the original ticket, but is willing to purchase a ticket if he lost money (Kahneman and Tversky, 1984). Second, "relative thinking" captures the essence of this behavior – that people think about relative magnitudes, while "mental accounting" is not intuitive in the same way. Finally, people can use mental accounts and still not be affected by relative price differences, suggesting that mental accounting and relative thinking are two separate things. For example, one could treat the calculator and the jacket in the classic experiment by Tversky and Kahneman (1981) as belonging to different mental accounts, but he can still treat a $5 discount on the two items in the same manner and be willing to make the same effort to save $5 regardless of the good's price. For these reasons, I use "relative thinking" to refer to this behavior in the rest of the article.

jacket's price was $15, only 29 percent were willing to drive.[2] Similar results were later obtained in a few additional experiments. Mowen and Mowen (1986) showed that not only students, but also business managers, exhibit this behavior. Ranyard and Abdel-Nabi (1993) varied the price of the second item (the jacket) and obtained similar results, and Frisch (1993) showed that the effect holds also when only a calculator is being purchased.[3] Azar (2006) showed that when subjects can purchase a certain good either in a store they currently visit or in a remote store, the minimal price difference for which they are willing to travel to the remote store is an increasing function of the price of the good they want to purchase. In an experiment that included 9 different price-treatments, he found that people behave (on average) as if the value of their time is approximately proportional to the square root of the good's price.

While the studies mentioned above focus on the trade-off between spending time and finding a cheaper price of a certain good, two recent studies show that a similar behavior exists when people consider differentiated goods or services and have a trade-off between the quality difference and the price difference. Once again, in scenarios in which only the absolute price difference should matter, people are affected also by the relative price difference. For example, Azar (2004) showed that consumers' willingness to add money for a high-quality good or service

---

[2] To see why such behavior does not fit standard economic theory, suppose, in line of the experimental results above, that one is willing to drive 20 minutes to save $5 on a certain cheap item, but then refuses to drive 20 minutes to save $8 on a more expensive item. This person could make the exact same effort of driving and yet be richer by $3 if he made the opposite choices, suggesting that his behavior is irrational.

[3] A few papers replicate these results and also find conditions that affect relative thinking, for example the range of percentages that could be saved (Darke and Freedman, 1993), the level of absolute savings (Moon, Keasey, and Duxbury, 1999), and reversing the scenario so that individuals trade-off money spent for time saved (Duxbury et al., 2005).

(over the price of a low-quality substitute good or service) is higher when the good's price is higher. The quality difference in his experiments was unrelated to the good's price and therefore the willingness to add should be independent of the good's price. The results were obtained both with undergraduate students and with economists participating in the 2003 North American Summer Meetings of the Econometric Society.

The literature discussed above suggests that there is a strong bias of relative thinking when people compare prices. An interesting question that was never explored before is whether people also exhibit relative thinking when they receive money rather than spend it. For example, when one receives a payment for a job he does, and he compares two possible payments (for example from two different employers, or the expected payment for low effort versus high effort), do the relative differences between the payments affect his decisions, just as relative price differences do? Intuitively, we can expect to observe a similar relative thinking effect. After all, compensation for work done is the price of labor, and if people exhibit relative thinking with respect to other prices, it seems reasonable that they will also do so with respect to the price of labor.

One important implication of relative thinking when receiving payments is for mixed compensation schemes that include both a fixed component and a variable component that depends on performance. Such compensation schemes are very common in real life. For example, salespeople often receive a base salary plus a percentage (or another function) of their sales. Similarly, many workers in the investment banking and consulting industries, as well as managers in different levels throughout the economy, receive a base salary plus a performance-based bonus. In other cases, managers receive a salary and also have options or stocks of their firm, so their compensation consists of their salary plus the return on these options or stocks, which in turn depends on their performance. It is therefore of great importance to examine

whether people exhibit relative thinking in the context of mixed compensation schemes. Better understanding of this issue, for example, can help firms choose the optimal mix of the fixed and variable components in the compensation schemes of millions of workers.

How does relative thinking relate to mixed compensation schemes? Relative thinking suggests that people consider relative magnitudes in addition to absolute magnitudes even when economic theory implies that only the latter should matter. When a person has to choose how much effort to exert in a certain task, relative thinking implies that he considers not only the absolute amount he can earn by exerting more effort, but also the relative increase in his earnings; that is, he might compare the additional earnings (due to higher effort) to his base salary (or to his total compensation). Consequently, a larger base salary may reduce effort because it makes the pay-for-performance bonus look smaller. This can have a dramatic effect on the optimal choice of incentive schemes by firms. Currently, the common wisdom is that increasing the fixed component of the salary can either have no effect on effort (because it does not depend on performance), or it may increase effort due to worker's reciprocity[4] or because of efficiency-wage arguments.[5] Relative thinking suggests that an opposite effect might also be present, and that if this effect is strong enough to outweigh the reciprocity and efficiency-wage effects (if these are present), then increasing the fixed payment to workers might in fact reduce effort.

---

[4] The idea is that the worker feels more grateful to his employer when he receives a higher salary, and because he wants to reciprocate, his effort is an increasing function of his salary (even though the salary is fixed and does not depend on performance). There is abundant experimental evidence for such behavior; for an excellent review of some of this experimental research, see Fehr and Gachter (2000).

[5] The efficiency-wage argument says that when a worker receives a higher salary, he becomes more afraid of losing his job (because the loss caused by losing the job and obtaining another job becomes higher), and therefore he makes more effort in order to reduce the risk that he will be fired.

The rest of the article is organized as follows. The next section presents the main experiment. Surprisingly, it turns out that the strong bias of relative thinking does not carry over from the domain of money paid (prices) to the domain of money received (compensation). The experiment shows that when people receive a fixed payment and a bonus for each task completed successfully, the relative magnitude of the bonus (compared to the fixed payment) does not affect their effort, suggesting that people do not exhibit relative thinking in this context. In order to identify whether relative thinking disappears (compared to hypothetical-questions experiments about price differences) because of the different context or because of the introduction of financial incentives, Section 3 presents another experiment that is almost identical to the first, but uses hypothetical questions without financial incentives. Relative thinking is not detected there either, suggesting that the different context, and not the introduction of financial incentives, is what eliminated relative thinking in the first experiment. The last section concludes.

2. The Incentive Condition Experiment

2.1. Experimental design

The main purpose of the experiment was to test whether in the context of task performance with mixed compensation schemes the bias of relative thinking exists. In order to create a decision problem in which the prediction of economic theory differs from that of relative thinking in a testable manner, the experiment involved a fixed payment and a variable payment. The variable payment was identical in all treatments – each correct answer increased the subject's earnings by 0.15 Shekels (1 Shekel was about $0.22 at the time the experiment was run). The fixed payment, however, differed in the two treatments. In the low-fee treatment, the subject

received a participation fee of 5 Shekels, while in the high-fee treatment he received 15 Shekels.[6] The treatment the subject received was randomized.

Because economic theory suggests that only incentives that depend on effort should affect the choice of effort, it follows that there should be no difference in effort in the two treatments. If relative thinking carries over from the domain of prices to the domain of payments for task performance, however, it suggests that the relative magnitude of incentives also plays a role in the choice of effort. When relative magnitudes are also considered, the same payment per correct answer seems larger when it is compared to a smaller fixed payment. This implies that subjects should make more effort in the low-fee treatment.

If the amounts involved were huge, economic theory could make a similar prediction because of wealth effects: when participation fee is higher the subject becomes richer, his marginal utility from money decreases, and therefore he makes less effort. But with the negligible amounts involved (the difference in the participation fee between the low-fee and high-fee treatments is slightly more than $2), it seems safe to assume that wealth effects can be ignored. Moreover, as we will see later, the results in fact indicate that subjects did not exert more effort in the low-fee treatment, and in addition, Part C of the experiment provides additional evidence that wealth effects do not seem to be an issue here.

---

[6] The range of possible fixed payments is limited for several reasons. A very low payment might create resentment and uncooperativeness of the subject. A payment too high might create an effect of willingness to stay in the experiment longer because of reciprocity issues (see the discussion below). The values of 5 and 15 Shekels were chosen so that they are in the reasonable range given the time the subject had to spend in the experiment, and yet to create a significant difference in the treatment in relative terms – one payment being 3 times higher than the other. It should be noted that in hypothetical questions about purchasing goods, a ratio of 1:3 in the price treatments is sufficient to produce a stark difference in responses (see for example Azar, 2004).

One potential confounding factor in the experiment could be reciprocity. Many studies show that people want to reciprocate and are willing to pay a price to do so (for an excellent review of some of this literature, see Fehr and Gachter, 2000). A subject who receives a higher participation fee might feel more willing to reciprocate, and if he thinks that the experimenter wants him to stay as much as possible in the experiment, this can lead to more effort in the high-fee treatment (in opposite direction to the possible effect of relative thinking). In order to avoid this potential confounding effect, two measures were taken.

The first measure was telling the subject several times, both in the consent form and in the instructions of Part B of the experiment (see Appendix A), that he can solve as many questions as he likes, he can quit anytime, and he is allowed not to solve the questions in Part B at all. This should eliminate the potential subject's belief that the experimenter wants him to stay in the experiment as much as possible and therefore that staying longer is a way to reciprocate.

The second measure was adding Part A at the beginning of the experiment. In this part, the subject was asked to answer 4 questions about various consumer decision problems. The participation fee of 5 or 15 Shekels was described as a payment for answering this part of the experiment. Consequently, after answering Part A subjects presumably felt that they completed the task for which they were paid (answering these four questions) and therefore they did not feel obligated to solve questions in Part B just in order to reciprocate and "pay with their time" for the participation fee.

In Part B the subjects were given the opportunity to solve up to 40 questions, where each correct answer increased their earnings by 0.15 Shekels. In order to give relative thinking its best chance, the earnings in all the parts of the experiment (Parts A, B, and the lottery in Part C if chosen) were paid together at the end. In addition, the earnings in Part B were described as a bonus in percentage of the amount earned in Part A (see Appendix A). It seems that if relative

thinking exists in the context of task performance, it should be stronger under these conditions than if each part was paid separately and the earnings in Part B were described as 0.15 Shekels for each correct answer.

The purpose of the questions was to measure effort, and therefore they were designed as a particularly boring task: answering which letter appears on a certain page, in a certain line and a certain location. Each subject was given 9 pages with letters (each subject received the same letters, but not the actual pages used before by another subject, to avoid the chance that a previous subject marked something on the pages). Three different measures of effort were employed. First, the number of attempted questions was recorded. Second, accuracy is presumably also positively correlated with effort, so the number of correct answers is a second measure, which takes also accuracy into account.

Finally, the research assistants also marked (without the subject paying attention to it) the time in which they gave the subject parts A and B (which were stapled together; the detailed description of the experimental procedure appears in Appendix A) and the time in which these parts were returned, thus allowing to compute the time the subject spent answering these parts. Because we are interested in the effort in Part B, this measure is a little noisy because it includes also the time it took the subject to answer part A. The variation in the time spent on Part A, however, is much smaller than the variation in time dedicated to Part B, because in Part A every subject answered the same questions, while in Part B some subjects did not answer any question while others solved up to 40 questions. Consequently, the amount of noise incorporated in this measure is not very large. Obviously, the three measures of effort are highly correlated; yet, looking at three different measures provides some additional robustness to the results.

To avoid the noise that can result from social influences among subjects (e.g., subjects continuing to solve questions as long as others do not leave, and stopping to solve when most

others left), the experiment was conducted individually and not in large groups. This design also allowed to give subjects their earnings as soon as they finished the experiment, without having to wait for others, thus creating a significant opportunity cost to solving more questions.[7] The cost of running the experiment individually was that it required many hours of research assistance, since the research assistants supervised the subjects during the experiment.

In order to avoid a situation where almost all subjects try to solve everything (or not to solve anything), the questions were presented in an increasing difficulty order (the font size became smaller and the pages became more congested as the subject proceeded in the questions). This design worked nicely and indeed there was significant heterogeneity in the number of questions attempted by different subjects (and therefore also in the number of correct answers and the time spent). The subjects were recruited on the campuses of two large Israeli universities, and were mostly undergraduate students in various academic fields. In total 227 subjects participated in this experiment; 118 in the low-fee treatment and 109 in the high-fee treatment (the numbers are unequal because treatment was randomized using a dice).

2.2. Results and discussion

Table 1 presents the distribution of the numbers of attempted questions and correct answers. We can see in Table 1 that on average subjects in the low-fee treatment attempted 1.2 questions more, and solved correctly 0.9 questions more, compared to subjects in the high-fee treatment. Given the large standard deviations of the numbers of questions attempted and solved correctly,

[7] If many subjects participate simultaneously with everyone being paid together in the end, there is no significant opportunity cost to solving more questions, and in fact it might be less boring to solve questions than to wait for others to finish. Such an alternative design therefore undermines the purpose of measuring effort by the number of questions attempted or answered correctly or by the time dedicated to the task.

however, it is clear that the difference between the treatments is not statistically significant, a point that can be observed also in the regressions reported in Table 2 and discussed below.

As mentioned earlier, a potential reason for lower effort of subjects in the high-fee treatment could be wealth effects together with the concavity of the utility function. It was explained that the amounts involved are too small for wealth effects to play a role, and moreover, the results reveal that there is no significant difference in effort between the two base-fee treatments. Nevertheless, as a robustness check (and because beforehand it was not known that there would be no difference in effort), a fair lottery was incorporated in Part C of the experiment. Subjects could accept or reject a fair lottery in which they had a probability of 1/6 to win 5 Shekels, and a probability of 5/6 to lose 1 Shekel. Every subject who has decreasing marginal utility from money even for changes of a few Shekels (which is the condition for wealth effects to be present in the choice of effort) should reject the lottery even if no transaction costs are involved (rejection should be even more common if transaction costs exist).[8] The result was that out of 227 subjects, 202 chose to accept the lottery. This is a surprisingly large fraction of acceptances, and it gives additional evidence that wealth effects do not seem to play a role in the decision of subjects about how many questions to solve.

Finally, in Part D subjects were asked to what extent the percentage increase and the absolute increase in earnings they could gain by solving more questions affected their decision when to stop solving the questions. We can mark the answer to the first question by p (importance of

[8] Notice that the reverse is not true: a subject who rejects the lottery does not necessarily do so because his marginal utility from money is decreasing for changes of a few Shekels. He might reject the lottery but actually be indifferent whether to accept it or not, or he might reject it because of transaction costs (spending more time in the experiment).

percentage increase) and to the second question by a (importance of absolute increase) and define a variable that measures relative thinking on a 0-1 scale as follows:[9]

$$R = [1.25p/(p + a)] - 0.125 \ .$$

The variable R satisfies several desired properties. It increases in p and decreases in a; it is equal to 1 in the extreme case that $p = 9$ and $a = 1$ and equal to 0 in the other extreme case (when $p = 1$ and $a = 9$); and it is equal to 0.5 when $p = a$. Thus, R can be interpreted as a measure of relative thinking expressed by the subject in the debriefing part (Part D). If there is a relationship between what people actually do and how they interpret their thoughts and actions in the debriefing part, then we should expect that the negative effect of high base fee on effort (if it exists) will be stronger for people with higher values of R.

Table 2 presents the results of several regressions. For reasons that will become clear shortly, this experiment is called the "incentive condition." The dependent variable is one of the three measures of effort: CORRECT (the number of correct answers), ATTEMPT (the number of questions attempted), and TIME (the time in minutes the subject spent on parts A and B together). The independent variables include the dummy variables MALE (equal to 1 for males), HIGHBASE (equal to 1 in the treatment with the high base fee), and BGU (equal to 1 for subjects in Ben-Gurion University of the Negev and 0 for subjects in Tel Aviv University). Additional independent variables are R, which measures the extent of relative thinking reflected in the debriefing questions as explained above, and a few interaction terms.

The results, reported in the three left columns of Table 2, show that with all three dependent variables, the effect of the base fee is small and statistically insignificant, including the

[9] The sample means of p, a and R are 4.30, 3.83, and 0.54, and their standard deviations 3.03, 2.99, and 0.20, respectively.

interaction terms of HIGHBASE with R and MALE. This implies that the relative magnitude of the payment for each correct answer being different in the two treatments did not affect the effort subjects made to solve questions. In other words, no relative thinking was documented in this experiment. In addition, the interaction term between HIGHBASE and R is positive, although it is not statistically significant. Notice that if answers to Part D were consistent with behavior, a higher value of R would reflect more relative thinking, which in turn implies that this interaction term should be negative (the more relative thinking a subject exhibits, the larger the negative effect of high base-fee on effort). This suggests that asking people whether they thought about absolute or relative differences may be uninformative in measuring the extent of relative thinking of different people.

The results are intriguing. As mentioned in the introduction, many experiments show that people exhibit strong relative thinking when answering hypothetical questions about the effort they would make to find a lower price of a good they want to purchase, or about their willingness to pay for two differentiated goods. Surprisingly, however, we fail to detect here any relative thinking, despite giving it its best shot by integrating the payment on Parts A and B and by stating the bonus per correct answer as a percentage of the earnings in Part A rather than as an absolute amount. This seems to suggest that people do not exhibit relative thinking in the context of task performance with mixed compensation schemes. Before we can be confident in this conclusion, however, there is one other potential explanation we should rule out.

This other potential explanation is the introduction of financial incentives.[10] In psychology it is very common and acceptable to conduct experiments that lack financial incentives, but in

[10] "Financial incentives" here and below should be interpreted as incentives that depend on the subject's performance, not just a fixed participation fee.

economics some researchers are skeptic about results that come from experiments without financial incentives.[11] Since all the previous studies on relative thinking used hypothetical questions, while the experiment here included financial incentives, it might be the case that the introduction of financial incentives, and not the different context, is what eliminated relative thinking here.

3. The Hypothetical Condition Experiment

3.1. Experimental design

In order to determine whether relative thinking was not detected because of the different context or the introduction of financial incentives, another experiment was conducted by replicating the former experiment but without financial incentives. Subjects went through the same steps as in the first experiment, but instead of actually solving the questions for real money, they were asked hypothetically (after solving three questions of different difficulty levels to get a feeling of the task) how many questions they think they would choose to solve if they received payment for each correct answer (the payment described was identical to that used in the incentive condition experiment).

While the incentive condition experiment differs from the previous literature both in its context (task performance vs. price comparisons) and in providing financial incentives, in the hypothetical condition only the context is different, because like the previous literature, this

[11] Several articles discuss the issue of financial incentives and how they affect experimental results, see for example Jenkins et al. (1998), Camerer and Hogarth (1999), Hertwig and Ortmann (2001a, 2001b, 2003) and Harrison and List (2004). The conclusions are generally that in some cases financial incentives affect behavior and choices, while in other cases they do not.

experiment also lacks financial incentives. Consequently, the results can suggest what eliminates relative thinking in the first experiment. If relative thinking is detected in the hypothetical condition, while it was not documented in the incentive condition, this will suggest that providing financial incentives eliminates the relative thinking bias. If relative thinking is not detected here, this will imply that the different context is what eliminates relative thinking, and that people do not exhibit relative thinking in the context of task performance with mixed compensation schemes.

The subject pool of the hypothetical condition experiment was similar to that in the incentive condition, and consisted of 89 students in the same two universities as before. Appendix B includes the wording of the experiment in this condition when it differs from that of the incentive condition. Three different measures of effort were used: how many questions the subject believed he would attempt, how many questions he thought he would solve correctly, and how much time he thought it would take him. Notice that this last measure is not directly comparable to that in the incentive condition, because there the time recorded includes the time spent on answering Part A and reading the instructions in Part B, and here the question refers only to how much time solving the questions in Part B would take. This does not prevent us from using this effort measure, however, because the goal is to compare the responses between the low-fee and high-fee treatments, and not to measure whether subjects correctly evaluate how much time it will take them to solve questions.

3.2. Results and discussion

Table 3 presents the results of the t-tests for difference in means and the Mann-Whitney tests that examine whether the three effort measures are different between the low-fee and the high-fee treatments. It is easy to see that the means are almost identical in the two treatments and that none

of the tests performed shows a statistically significant difference between the two treatments, indicating that subjects exhibit no relative thinking. To perform yet another test of relative thinking, the data from the two experiments were combined, and a dummy variable HYPOTH which equals 1 in the hypothetical condition and 0 in the incentive condition was defined. Running the three previous regressions with the addition of the independent variable HYPOTH and the interaction between HIGHBASE and HYPOTH on the combined sample of the two conditions shows that the coefficient of this interaction is very close to zero and is statistically insignificant (see the right three columns of Table 2). If relative thinking existed in task performance and financial incentives were the reason it disappeared in the incentive condition, this interaction coefficient should be negative.

This and the tests presented in Table 3 suggest that subjects in the hypothetical condition did not exhibit relative thinking. Since this condition lacks financial incentives and its conclusion of no relative thinking is similar to the conclusion in the incentive condition, we can infer that what eliminates relative thinking in these experiments compared to the previous literature is the different context and not the introduction of financial incentives.

The statistically significant coefficient of HYPOTH when the dependent variable is TIME is not surprising, because in the incentive condition TIME includes the time it took the subject to answer Part A and to read the instructions of Part B, while in the hypothetical condition it only refers to how much time the subject thought it would take him to solve the questions in Part B. Other than HYPOTH, none of the independent variables is statistically significant at the 5% level or below.

The two conditions also allow us to examine a couple of other questions not related to relative thinking, that deal with the differences between the behavior people think they will choose and what they actually choose. The separation between the two conditions is an advantage: if the

same person is first asked what he thinks he will do (e.g., how many questions he will attempt solving) and then asked to actually do something (e.g., solve questions), his response in the first part can influence what he chooses to do in the second part (for example because he wants to be consistent with what he said earlier). Consequently, it will not be surprising in such an experiment to find that what people think they will do is similar to what they later do. Here, however, the people who answer the hypothetical questions are not the ones who actually make choices, so it is less obvious whether the answers in the two conditions are going to be similar. Because the characteristics of the subject pool in the two conditions are the same, if people are good at predicting their behavior (at least in contexts similar to the one in the experiment) then we should observe similar outcomes in the hypothetical questions and the actual behavior.

We can see in Table 2 that the coefficient of HYPOTH in the regression of ATTEMPT is negative but not statistically significant, indicating that people slightly underestimate the number of questions they will solve. A closer look at the data, however, reveals a striking difference in the percentage of people who do not solve any question: in the hypothetical condition, 18 out of 86 subjects (20.9%) who gave usable responses to that question indicated that they would not attempt to solve any questions. In the incentive condition, only 7 out of 227 subjects (3.1%) did not solve any question. A possible reason for this difference may be that the subjects in the incentive condition wanted at least to get a feeling of what the task requires, while the subjects in the hypothetical condition were asked anyway to solve three questions first in order that their responses for the hypothetical questions would be more meaningful (see Appendix B).

The coefficient of HYPOTH in the regression of CORRECT shows that subjects were quite accurate in their prediction of how many questions they would solve correctly. Notice, however, that CORRECT is affected both by how many questions the subject attempts and by how accurate he is. To examine accuracy only, we can consider the percentage of questions solved

correctly, CORRECT / ATTEMPT, as a measure of accuracy. Are people optimistic, thinking they will make less mistakes than they actually do? It turns out that this effect is very small and not statistically significant: average accuracy in the incentive condition is 88.4% and in the hypothetical condition it is 90.1%, and the p-value of the t-test for difference in means is 0.328.

4. Conclusion

A lot of experimental evidence shows that when people compare prices of the same good in different stores or prices of differentiated goods, the relative price differences affect their decisions even when economic theory implies that only absolute price differences should matter, a phenomenon that was recently denoted "relative thinking." This article presents an experiment that examines whether relative thinking also exists in the context of task performance with mixed compensation schemes that include both a fixed and a pay-for-performance components. The results show that relative thinking does not exist in this context: the ratio between the pay-for-performance compensation and the fixed compensation does not affect effort. An additional difference between this study and the previous ones, however, is that the experiment here included financial incentives, while the previous studies used hypothetical questions. This raises the question what is the reason relative thinking disappears here: is it the different context (payments received for task performance instead of price comparisons), or the introduction of financial incentives?

To address this question, another condition was run. The second condition was almost identical to the first, except that instead of making actual decisions about how much time to stay and how many questions to solve, the subjects were asked hypothetically about these decisions. No relative thinking was documented in this condition either. This suggests that in the context of

task performance with mixed compensation schemes there is no relative thinking, regardless of whether financial incentives are provided or not. This is an important finding given that it is the first examination of relative thinking in this context, and given the prevalence of mixed compensation schemes in the remuneration of managers, salespeople, and other workers in various industries.

Appendix A: Instructions in the Incentive Condition[12]

Procedure: First the subject received an informed-consent form to sign. After returning it, a dice was thrown and according to the outcome the base-fee treatment was chosen and the subject was handed the appropriate version of Part A and Part B stapled together and the 9 pages containing the letters on which the questions in Part B are based. When he returned them, he received Part C and Part D stapled together. After filling Part C, if he was interested, the lottery was conducted. At the end he received his payment in cash according to the amount he earned in all the parts of the experiment together.

Consent form

The purpose of this experiment is to examine various aspects of decision making. In the first part you will be asked to answer a few short questions. The second part is optional: you can end it and return the questionnaire anytime, and you are even allowed not to answer any question in this part. Afterwards you will be able to participate in a lottery if you want, and you will be asked to

[12] This is a translation of the Hebrew original. The brackets indicate the numbers used in the low base-fee treatment (the bonus percentage is higher when the base fee is lower because the payment for each correct answer is identical in both treatments).

answer two additional short questions. At the end you will receive the amount that you earned in the experiment in cash.

The questionnaire is anonymous. Participation in the experiment is voluntary and you can leave the experiment anytime.

Participant's declaration: I agree to participate in this experiment.

Signature and date:

Part A

Thank you very much for your participation in the experiment. The questions in this part have no right or wrong answer. Try to imagine yourself in the situations described in the questions before you answer. For your participation in this part of the experiment you will receive 15 [5] Shekels at the end of the experiment.

[The subject was given 4 questions in which he was asked to provide matching prices in various consumer decision problems; the questions are omitted for the sake of brevity and are available from the author upon request].

Part B

This part is optional. You can end it and return the questionnaire any time, even if you chose not to answer any question. For your participation in this part you can earn an additional bonus. If you solve all 40 questions in this part correctly, you will get a bonus of 40% [120%] (in relation to the amount you earned for part A, i.e., you can receive in total for the two parts 1.4 [2.2] times the payment for part A). If you solve some of the questions correctly, you will receive a pro-rated bonus according to the number of questions you solve correctly. For example, if you solve 20 questions correctly, you will receive a bonus of 20% [60%]. As you will see, the questions

become more and more difficult, so it is recommended to proceed according to their order, up to the point in which you decide that you do not want to continue solving additional questions.

On the following pages different letters are written. In the questions below you are asked to write which letter appears on a certain page, in a certain line, and in a certain location. The location of the letter in the line is counted from left to right. For example, the letter on page 1, line 1, column 3 is s.

1. Page 1, line 1, location 2:_____
2. Page 1, line 5, location 6:_____

…

40. Page 9, line 31, location 68:_____

Part C – Lottery (Optional)

You can now participate in the following lottery: you guess a number between 1 and 6, and we throw a dice. The participation in the lottery will cost you 1 Shekel (which will be reduced from your earnings in parts A and B), but if the number you guessed is equal to the number that will be on the dice, you will earn 6 Shekels (i.e., together with the cost of the lottery, you have in fact a probability of 5/6 to lose 1 Shekel and a probability of 1/6 to earn 5 Shekels). If you want to participate in the lottery, please write the number you guess: ____

Part D

In order to compute the amount you earned, I need to check your responses in part B. I will be grateful if you can answer in the meantime the following questions (they do not have a right or wrong answer, simply write whatever is valid for you):

Rank on a scale of 1 to 9 (circle the answer you choose) to what extent each of the following reasons affected your decision when to stop solving the questions in Part B (1: this reason was not relevant at all; 9: this reason was my main consideration).

I compared the time and effort required to solve the questions to the fact that I could increase my earnings by 40% [120%] by solving additional questions: 1 2 3 4 5 6 7 8 9

I compared the time and effort required to solve the questions to the fact that I could increase my earnings by 6 Shekels by solving additional questions: 1 2 3 4 5 6 7 8 9

Please write below additional comments, if you have, about the experiment and the manner in which you made decisions in it:

[Space provided but omitted here]

Thank you very much for your participation in the experiment!

Appendix B: Instructions in the Hypothetical Condition[13]

Procedures, Consent form, and Part A: identical to the incentive condition version.

Part B

On the following pages different letters are written. In the questions below you are asked to write which letter appears on a certain page, in a certain line, and in a certain location. The location of the letter in the line is counted from left to right. For example, the letter on page 1, line 1, column 3 is s.

[13] This is a translation of the Hebrew original. The brackets indicate the numbers used in the low base-fee treatment.

First, you are asked to solve the three questions 10, 20 and 30 in order to get a feeling about what is required in the questions. After you finish these three questions, please answer the following questions:

Suppose that you were told the following:

> This part is optional. You can end it and return the questionnaire any time, even if you chose not to answer any question. For your participation in this part you can earn an additional bonus. If you solve all 40 questions in this part correctly, you will get a bonus of 40% [120%] (in relation to the amount you earned for part A, i.e., you can receive in total for the two parts 1.4 [2.2] times the payment for part A). If you solve some of the questions correctly, you will receive a pro-rated bonus according to the number of questions you solve correctly. For example, if you solve 20 questions correctly, you will receive a bonus of 20% [60%]. As you will see, the questions become more and more difficult, so it is recommended to proceed according to their order, up to the point in which you decide that you do not want to continue solving additional questions.

How many questions out of the 40 questions below do you think you would have tried to solve? _____

How many of them do you think you would have solved correctly? _____

How much time do you think it would have taken you to solve the number of questions you indicated? _____

Comment: The description above is hypothetical only, please do not solve the 40 questions below.

1. Page 1, line 1, location 2:_____
2. Page 1, line 5, location 6:_____

…

40. Page 9, line 31, location 68:_____

Part C: identical to the incentive condition version.

Part D

I will be grateful if you can answer the following questions (they do not have a right or wrong answer, simply write whatever is valid for you):

Rank on a scale of 1 to 9 (circle the answer you choose) to what extent each of the following reasons affected your decision about how many questions you would choose to solve in Part B (1: this reason was not relevant at all; 9: this reason was my main consideration).

[The rest of this part is identical to the incentive condition version.]

References

Azar, Ofer H. (2004): "Do People Think about Absolute or Relative Price Differences? Experimental Evidence and Implications," Working paper, Ben-Gurion University of the Negev.

Azar, Ofer H. (2006): "Are Transportation and Search Costs Independent of the Good's Price? Some Experimental Evidence," Working paper, Ben-Gurion University of the Negev.

Camerer, Colin F. and Robin M. Hogarth (1999): "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty* 19, 7-42.

Darke, Peter R. and Jonathan L. Freedman (1993): “Deciding Whether to Seek a Bargain: Effects of Both Amount and Percentage Off,” *Journal of Applied Psychology* 78:6, 960-965.

Duxbury, Darren, Kevin Keasey, Hao Zhang and Shue Loong Chow (2005): “Mental Accounting and Decision Making: Evidence Under Reverse Conditions Where Money Is Spent for Time Saved,” *Journal of Economic Psychology* 26, 567-580.

Fehr, Ernst and Simon Gachter (2000): "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14:3, 159-181.

Frisch, Deborah (1993): “Reasons for Framing Effects,” *Organizational Behavior and Human Decision Processes* 54, 399-429.

Harrison, Glenn W. and John A. List (2004): "Field Experiments," *Journal of Economic Literature* 42, 1009-1055.

Hertwig, Ralph and Andreas Ortmann (2001a): “Experimental Practices in Economics: A Challenge for Psychologists?” *Behavioral and Brain Sciences* 24, 383-403.

Hertwig, Ralph and Andreas Ortmann (2001b), “Money, Lies, and Replicability: On the Need for Empirically Grounded Experimental Practices and Interdisciplinary Discourse," *Behavioral and Brain Sciences* 24, 433-451.

Hertwig, Ralph and Andreas Ortmann (2003): “Economists’ and Psychologists’ Experimental Practices: How They Differ, Why They Differ, and How They Could Converge,” in Brocas, Isabelle and Juan D. Carrillo (Eds.), *The Psychology of Economic Decisions, Volume I: Rationality and Well-Being*, Oxford University Press, Oxford, pp. 253-272.

Jenkins, G. Douglas, Jr., Atul Mitra, Nina Gupta and Jason D. Shaw (1998): "Are Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research," *Journal of Applied Psychology* 83, 777-787.

Kahneman, Daniel and Amos Tversky (1984): "Choices, Values and Frames," *American Psychologist* 39, 341-350.

Moon, Philip, Kevin Keasey and Darren Duxbury (1999): "Mental Accounting and Decision Making: The Relationship between Relative and Absolute Savings," *Journal of Economic Behavior and Organization* 38, 145-153.

Mowen, Maryanne M. and John C. Mowen (1986): "An Empirical Examination of the Biasing Effects of Framing on Business Decisions," *Decision Sciences* 17, 596-602.

Ranyard, Rob and Deborah Abdel-Nabi (1993): "Mental Accounting and the Process of Multiattribute choice," *Acta Psychologica* 84, 161-177.

Thaler, Richard H. (1985): "Mental Accounting and Consumer Choice," *Marketing Science* 4, 199-214.

Thaler, Richard H. (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making* 12, 183-206.

Tversky, Amos and Daniel Kahneman (1981): "The Framing of Decisions and the Psychology of Choice," *Science* 211, 453-458.

Table 1: Attempted Questions and Correct Answers in the Incentive Condition

	Attempted questions		Correct answers	
	Low fee	**High fee**	**Low fee**	**High fee**
0-10 questions	30 (25.4%)	32 (29.4%)	36 (30.5%)	37 (33.9%)
11-20 questions	32 (27.1%)	30 (27.5%)	31 (26.3%)	30 (27.5%)
21-30 questions	12 (10.2%)	12 (11.0%)	14 (11.9%)	12 (11.0%)
31-40 questions	44 (37.3%)	35 (32.1%)	37 (31.4%)	30 (27.5%)
Total	118 (100%)	109 (100%)	118 (100%)	109 (100%)
Average	23.3	22.1	20.6	19.7
Standard deviation	13.8	13.6	12.6	12.5

The numbers in the top rows indicate the number of subjects in each cell. The numbers in the bottom two rows are the average and standard deviation of the number of attempted questions or correct answers in each treatment.

Table 2: Regression Results

Condition	Incentive	Incentive	Incentive	Incentive + hypothetical	Incentive + hypothetical	Incentive + hypothetical
Dependent var. / Independent var.	CORRECT	ATTEMPT	TIME	CORRECT	ATTEMPT	TIME
Constant	22.51**	23.90**	23.75**	25.96**	28.28**	26.85**
	(3.97)	(4.29)	(4.26)	(3.57)	(3.82)	(3.53)
HIGHBASE	-1.15	-0.45	-0.43	-3.97	-4.51	-3.38
	(5.45)	(5.90)	(5.86)	(4.74)	(5.10)	(4.74)
R	-2.32	-0.49	-2.40	-9.50	-9.13	-6.58
	(6.28)	(6.79)	(6.75)	(5.59)	(5.95)	(5.47)
MALE	-2.16	-2.17	1.03	-1.42	-1.70	0.11
	(2.37)	(2.56)	(2.54)	(2.17)	(2.34)	(2.18)
BGU	1.22	1.90	2.88	1.00	1.62	2.21
	(1.77)	(1.91)	(1.90)	(1.58)	(1.70)	(1.58)
HIGHBASE*R	2.50	0.07	1.86	8.70	8.70	7.79
	(8.74)	(9.44)	(9.38)	(7.52)	(8.09)	(7.49)
HIGHBASE*MALE	-2.45	-2.08	-5.21	-3.12	-2.85	-5.64
	(3.42)	(3.70)	(3.68)	(3.03)	(3.26)	(3.04)
HYPOTH				-1.14	-2.65	-9.86**
				(2.63)	(2.80)	(2.59)
HIGHBASE *HYPOTH				-0.07	0.08	0.74
				(3.52)	(3.75)	(3.48)
N	221	221	221	303	307	308
R^2	0.03	0.03	0.04	0.03	0.04	0.14

Standard errors are reported in parentheses. The number of observations in the incentive condition is 221 because 6 subjects did not answer the questions in part D and therefore are omitted from the regressions. Levels of statistical significance are denoted by asterisks: * represents 5%-level and ** represents 1%-level. Notice that the average effect of a variable should also account for the interaction terms involving this variable, taking into account the average value in the sample of the interacting variables (in the incentive condition, the relevant means are as follows: MALE: 0.57, HIGHBASE: 0.48, R: 0.54. In the combined sample, the means are: MALE: 0.55, HIGHBASE: 0.51, R: 0.56, HYPOTH: 0.29. The means in the combined sample include a few observations that are not included in one or more regressions because of unusable answers to one of the effort measures).

Table 3: Hypothetical Condition Results

	Attempted questions	Correct answers	Time
Standard deviation	15.8	14.5	12.8
Low-fee mean	20	18.9	14.0
High-fee mean	19.6	18.6	13.6
p-value of the t-test for difference in means	0.913	0.946	0.899
Low-fee median	20	17	17.5
High-fee median	20	15	10
p-value of the Mann-Whitney test	0.936	0.955	0.453
N	86	82	87

The number of observations is slightly different among the three effort measures because of unusable responses.