# How financial incentives and cognitive abilities affect task performance in laboratory settings: an illustration

**Ondrej Rydval, Andreas Ortmann**[*]

*CERGE-EI, Prague, Czech Republic*

April 2004

**Abstract**

Drawing on Gneezy and Rustichini (2000), we show that cognitive abilities seem at least twice as important as financial incentives. This result backs up the exhortation of Camerer and Hogarth (1999) to pay attention to both the labor and capital aspects of cognitive production in experiments.

*Keywords:* Financial incentives, Cognitive abilities, Experiments, Field experiments

*JEL classification:* C81; C91; C93; D83

---

[*] Corresponding author: Andreas Ortmann, Center for Economic Research and Graduate Education, Charles University, and Economics Institute, Academy of Sciences of the Czech Republic (CERGE-EI), Address: Politickych veznu 7, 11121 Prague, Czech Republic, Tel: +420-224005117, Fax: +420-224227143, E-mail address: andreas.ortmann@cerge-ei.cz

## I. Introduction

One of the distinguishing features of the practices of experimental economists is performance-dependent subject payments, or financial incentives (Hertwig and Ortmann, 2001). In fact, among economists financial incentives have become a strictly enforced convention on the widely shared belief that decisions have to matter to those participating in experiments for the data to have meaning. This belief found its expression in Smith's nonsatiation and saliency tenets of proper experimentation (e.g., Smith, 1976, 1982) and, by and large, seems to be supported by empirical evidence produced by economists: In their survey article, Smith and Walker (1993) suggest that "increased financial rewards [may] shift the central tendency of the data toward the predictions of rational models…[and] in virtually all cases rewards reduce the variance of the data around the predicted outcome." (Smith and Walker, 1993, p. 245; see Hertwig and Ortmann, 2001, for a more differentiated assessment.) Smith and Walker's conclusion is generally referred to as the labor theory of cognition.

Two recent papers have urged a reconsideration of this view. Gneezy and Rustichini (2000) demonstrated in a thought-provoking experiment that increasing financial incentives does not monotonically lead to more favorable evidence for economic theories. Rather, these authors' results suggest – as succinctly summarized by the title of their article – that the experimenter ought to pay enough, or not pay at all. In other words, while economists' belief in the effects of increasing financial incentives in experiments seems to be right on the money for reasonably high financial incentives, microscopic payments have – for reasons not yet agreed on – detrimental effects on subject behavior. Gneezy and Rustichini (2000) conjecture that their minimally paid subjects might have

been insulted by the microscopic compensation offered to them and consequently performed worse than subjects who, apart from a flat participation fee, solved the tasks solely based on their intrinsic motivation.

Camerer and Hogarth (1999), importantly and in our view correctly, take a broader view and argue that the real problem lies in economists' focus on the labor aspect, and almost complete neglect of the capital aspect, of cognitive production in experiments. In their view, cognitive performance is affected not only by incentives and thus effort that subjects exert, but also by cognitive abilities that are fixed in the short run of the experiment and can be enhanced if learning is allowed. In turn, the authors argue, interpreting experimental results should be conditional not only on the particular financial incentives employed but also on cognitive abilities of the participants in the experiment.

And yet, while the arguments in Camerer and Hogarth (1999) are persuasive, they are informal. In fact, to the best of our knowledge there exists no persuasive empirical evidence produced by economists directly illuminating the *relative* importance of individual abilities and incentives in cognitive production[1], although there are some

---

[1] Palacios-Huerta (2003) examines the impact of tournament-type, and increased, incentives in the last 15 rounds of a repeated Monty Hall Three Door problem while controlling for subjects' GPA, learning, and several other treatment effects. The author finds that "more able individuals significantly respond to the size of incentives" (p. 247). However, this result is quite possibly due to the five sizeable cash prizes that induce aspects of a tournament-type competition rather than the increased incentives. The author also does not, and actually cannot because of the confounds introduced by the cash prizes, address the relative importance of individual abilities and incentives. In addition, it is not clear to what extent the impact of ability and incentives is confounded with social interactions allowed in the last 15 rounds. Eckel (1999) uses natural

precursors in related disciplines that are suggestive.[2] Here we provide such empirical evidence by drawing on the data in Gneezy and Rustichini (2000). Specifically, we will show that the effects of financial incentives seem much less important, even under the best of circumstances, than those of cognitive abilities approximated by a measure of subjects' IQ.

## II. The data

In Figure 1 below, we adapted individual-level data from Gneezy and Rustichini (2000) who examined the impact of financial incentives of different strength on performance in a psychometric (IQ) test. Importantly, the authors did not analyze their data at a disaggregate level. Such disaggregation provides us, however, with important quantitative insights about the separate effects of incentives and individual abilities on cognitive performance.

---

framing of lottery choices, which she argues is a substitute for increased incentives, and finds that individual performance is positively correlated with subjects' GPA score. She did not, however, run the obvious control treatment with increased incentives that would have completed her argument.

[2] In a unique example of interacting financial incentives and intrinsic ability, Awasthi and Pratt (1990) examined their relative impact in three accounting tasks. Contrary to what seems to hold for the data of Gneezy and Rustichini analyzed below, the authors find that incentive effects are significantly stronger for subjects with higher cognitive abilities (measured by the EFT test of perceptual differentiation, a very idiosyncratic measure of cognitive capital). However, the authors cannot do full justice to the estimation due to having only a binary performance measure. Therefore they are not able to quantify the relative importance of individual ability and incentives.

Gneezy and Rustichini randomly assigned 160 subjects to four treatments (no-pay, NIS0.1, NIS1 and NIS3; from here on "incentive treatments") and then examined the impact of financial incentives on average IQ score for each treatment.[3] They reported a non-monotonic impact of incentives on performance: average performance was highest and almost identical for the two high-incentive treatments (NIS1 and NIS3), but lowest for the NIS0.1 treatment. In fact, average performance in this treatment was statistically lower than performance in the no-pay treatment.

-- Figure 1 about here --

In Figure 1 we assume (as Gneezy and Rustichini did) that the subjects in the four incentive treatments were sampled from a common population, and we plot, for each of the four incentive treatments, individual IQ performance, in ascending IQ rank order. An individual's IQ performance, measured on the vertical axis, induces her or his IQ rank, indicated on the horizontal axis. Note that the individuals ranked "1" in each treatment scored worst, while those ranked "40" scored best. For each incentive treatment, the connected IQ-score observations yield a "performance curve" for that treatment. Whereas such a performance curve visually describes the *within-treatment* variation in performance, one can similarly inspect the *across-treatment* variation by making comparisons among the performance curves. Three observations are noteworthy:

---

[3] Subjects were also paid a flat participation fee of NIS60. At the time of the experiment, the exchange rate was 3.5 NIS (New Israeli Shekel) to $1. The fifty IQ-type questions, taken from a test normally used to screen university applicants, involved mainly reasoning and computation skills. The subjects were volunteer male and female undergraduate students at the University of Haifa from all fields of study with an average age of 23 years.

First of all, notice that the performance curves for the high-incentive treatments (NIS1 and NIS3) are virtually identical and slope considerably upwards, implying that there is a high within-treatment variation in performance but hardly any across-treatment one. Arguably, this is most likely due to a significant within-treatment variation in cognitive abilities. One could conceive that the large within-treatment performance variation is partly also effort-driven, but the variation in cognitive effort required to generate this result is unlikely; plus one would need to explain why the two performance curves seem almost identical despite the across-treatment incentive (and thus presumably effort) differential. Therefore, consistent with the interpretation of the IQ score as ability rank, it seems quite plausible that ability rather than incentive differentials determine individual performance differentials when incentives are high enough.

Next inspect the performance curves for the low-incentive treatments (no-pay and NIS0.1). Clearly, Gneezy and Rustichini (2000) were right in asserting that the NIS0.1 subjects overall were less motivated than the ones in the no-pay treatment. This is particularly apparent at the low-performance end where the gap between the performance curves for the low-incentive treatments widens (and, in addition, so does the gap between the performance curves of the two low-incentive treatments and the two high-inventive treatments). That the NIS0.1 subjects were less motivated than the ones in the no-pay treatment also seems confirmed by the performance curve for the NIS0.1 treatment lying below that for the no-pay treatment across the whole performance range. It is highly

unlikely that this would be caused by across-treatment ability differentials, and thus across-treatment differences in motivation must have played the main role.[4]

Finally and most importantly, focus on the slope of all four performance curves and the distance between them. An eyeball test reveals that, leaving aside the motivational problems at the low-performance end, the within-treatment variation in performance is generally much greater than the variation across treatments. To give a meaningful comparison, consider the largest across-treatment performance differential at the median rank. This turns out to be 13 (i.e., 24 correct answers in the NIS0.1 treatment vs. 37 in the NIS1 treatment), which is equivalent to the performance differential associated with moving up from the first to the third quartile within the NIS1 treatment (28 vs. 41). Note, however, that within-treatment performance differentials can be much larger. For instance, in both of the high-incentive treatments (NIS1 & NIS3), the difference in performance for individuals ranked 1 and 40 is as large as 34.

Provided that the across-treatment performance variation can be assigned to incentive effects while the within-treatment performance variation to ability differentials, an

---

[4] Clearly, effort is a function of motivation. But what we see at the low-performance end of the two low-incentive treatments is not only a lack of motivation to perform but an outright refusal to perform, and in any case a different form of "lack" of motivation. Gneezy (2004) proposes several possible explanations, such as participants feeling insulted or being at odds with their self-perception. That such reactions are much less prevalent in the no-pay treatment and entirely absent in the two high-incentive treatments (NIS1 and NIS3) seems to confirm the presence of a specific kind of motivational problem in the NIS0.1 treatment. It also seems to justify the current practice of most experimental economists of paying their subjects at least twice the minimum wage.

important and powerful result follows: ability differentials among individuals seem to account for a much greater part of performance variation than incentive effects.

**III. Discussion**

As to the motivational complications in the first performance quartile, we do not know whether the "unmotivated" subjects at the low-performance end of the no-pay and NIS0.1 treatments have low abilities: it is possible, although in our view implausible, that these subjects are high-ability individuals.[5] But this makes the case for the relative importance of cognitive-capital effects even stronger: were it not for motivational problems of (possibly) high-ability individuals, cognitive-capital effects could be even larger.

Our illustration backs up the exhortation of Camerer and Hogarth (1999) to pay attention to both the labor and capital aspects of cognitive production in experiments. To the extent that our illustration is representative of the relative importance of cognitive capital effects, not controlling for cognitive abilities is an important shortcoming of almost all existing experimental studies examining incentive effects.

Our illustration makes a strong case for considering cognitive abilities, as well as motivational factors, as major determinants of cognitive performance. This was clearly a major deficiency of the labor theory of cognition, one which needs to be remedied by integrating the insights of Camerer and Hogarth (1999), and of Gneezy and Rustichini

---

[5] In our view, low-ability participants have to exert more effort than high-ability ones, so we are unlikely to find the low-ability participants in the upper part of the performance curve. Reinforcing this argument, high-ability participants potentially face significantly higher reputational costs than their low-ability counterparts (Benabou and Tirole, 2003).

(2000), into that framework. For a very preliminary attempt at such a capital-labor framework, see Wilcox (1993) who proposes that to solve mental tasks, subjects employ algorithms of various sophistication and effort cost, yet he does not go into detail of what such algorithms consist of.

**References**

Awasthi, V. and J. Pratt, 1990, The effects of monetary incentives on effort and decision performance: the role of cognitive characteristics, Accounting Review 65, 797-811.

Benabou, R. and J. Tirole, 2003, Incentives and Prosocial Behavior, Manuscript.

Camerer, C. and R. Hogarth, 1999, The effects of financial incentives in experiments: a review and capital-labor-production framework, Journal of Risk and Uncertainty 19, 7-42.

Eckel, C., 1999, Comment on 'The effects of financial incentives in experiments: a review and capital-labor-production framework', Journal of Risk and Uncertainty 19, 47-48.

Gneezy, U., 2004, The W effect of incentives, Manuscript.

Gneezy, U. and A. Rustichini, 2000, Pay enough or don't pay at all, Quarterly Journal of Economics 115, 791-811.

Hertwig, R. and A. Ortmann, 2001, Experimental practices in economics: a methodological challenge for psychologists?, Behavioral and Brain Sciences 24, 383-451.

Palacios-Huerta, I., 2003, Learning to open Monty Hall's doors, Experimental Economics 6, 235-251.

Smith, V. L., 1976, Experimental economics: induced value theory, American Economic Review Proceedings 66, 247-279.

Smith, V. L., 1982, Microeconomic systems as an experimental science, American Economic Review 72, 923-955.

Smith, V. L. and J. Walker 1993, Monetary rewards and decision cost in experimental economics, Economic Inquiry 31, 245-261.

Wilcox, N. 1993, Lottery choice: incentives, complexity, and decision time, The Economic Journal 103, 1397-1417.
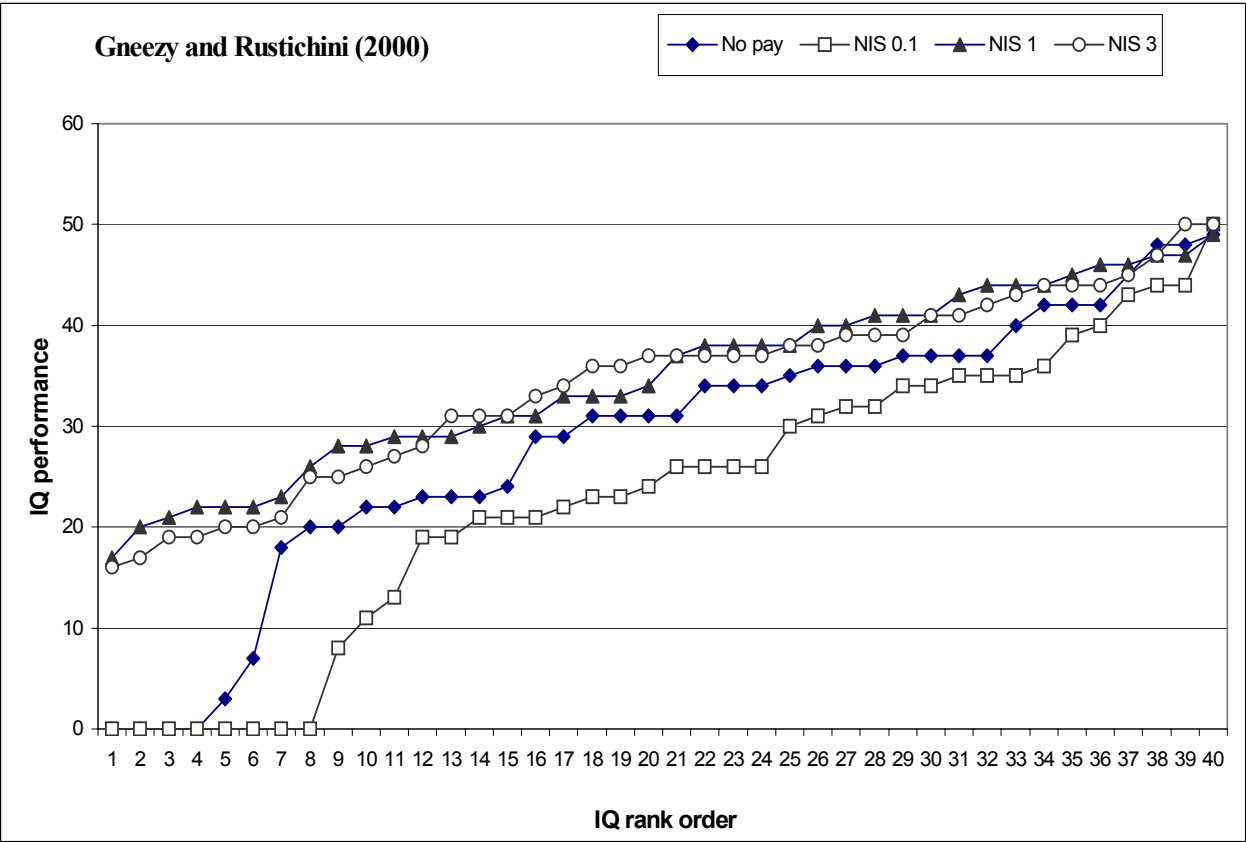
Figure 1: Individual IQ performance, plotted in ascending IQ rank order, separately for each of the four incentive treatments.