How financial incentives and cognitive abilities affect task performance in laboratory settings: an illustration[†]

Ondrej Rydval, Andreas Ortmann*

CERGE-EI

Abstract

Drawing on Gneezy and Rustichini (2000), we illustrate that subjects' cognitive abilities seem at least as important for their performance as do financial incentives they face. Theorists should thus pay more attention to the ability aspect of cognitive production.

Abstrakt

Na základě analýzy Gneezy and Rustichini (2000) ukazujeme, že kognitivní schopnosti subjektů mohou být minimálně tak důležité pro jejich výkonnost jako finanční incentivy, kterým čelí. Teorie by proto měla klást větší důraz na lidské schopnosti při modelování kognitivní činnosti.

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

JEL classification: C81; C91; C93; D83

[†] We thank D. Engelmann, U. Gneezy, R. Hogarth, and S. Jurajda for very helpful comments.

* Corresponding author: 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: +420-224-005-117, E-mail address: andreas.ortmann@cerge-ei.cz

I. Introduction

Subject behavior in economic experiments is often sensitive to the (performance-dependent) reward scheme and to the subject pool employed (e.g., Camerer, 2003; Hertwig and Ortmann, 2001). Sensitivity to financial incentives and subject-pool characteristics is an important methodological concern with implications for internal as well as external validity of experimental findings. Camerer and Hogarth (1999) were the first to informally propose a sensible link between financial incentives, subjects' cognitive characteristics, and their performance. These authors argued that a subject's performance depends on a combination of her mental effort (stimulated by her intrinsic motivation and financial incentives) and her task-relevant cognitive abilities (fixed in the short run of the experiment but possibly improvable through learning, if allowed).

While the arguments of Camerer and Hogarth are persuasive, there exists no satisfactory study addressing the *relative* importance of financial incentives and individual cognitive abilities for task performance. Awasthi and Pratt (1990) report that higher financial incentives are associated with significantly higher success rate in applying conjunction-probability and sample-size decision rules, but only for subjects with superior perceptual differentiation (as measured by EFT test score). A similar result is reported by Palacios-Huerta (2003), who found that more able individuals (in terms of GPA, SAT score, and high school rank) performed significantly better in his repeated Monty Hall Three Door problem and responded more strongly to higher incentives. Assessing the impact of incentives and abilities on performance based on these studies is nevertheless problematic. Awasthi and Pratt employed a very idiosyncratic measure of intrinsic cognitive abilities. Palacios-Huerta increased incentives concurrently with introducing tournament-type competition; another confound arose through allowing social interaction among participants. Furthermore, both studies only used binary measure of individual performance, hence precluding an accurate assessment of the relative impact of incentives and abilities thereupon.

We provide such an assessment drawing on the data in Gneezy and Rustichini (2000). Specifically, we show that the impact of financial incentives on individual performance seems less important, even under the best of circumstances, than that of cognitive abilities approximated by a measure of subjects' IQ.

II. The data

We adapted individual-level data from Gneezy and Rustichini (2000) who examined the impact of varying financial incentives on performance in a psychometric (IQ) test. Their 160 subjects answered 50 IQ-type questions involving mainly reasoning and computation skills similar to those found in the GMAT. In addition to being paid NIS60 for participating, subjects earned zero, NIS0.1, NIS1, or NIS3 for each correct answer depending on which of the four incentive treatments they were assigned to.[1]

Inspecting the average IQ score for each treatment, Gneezy and Rustichini 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 – significantly lower than in the no-pay treatment. This result

[1] At the time of the experiment, the exchange rate was 3.5 NIS (New Israeli Shekel) to $1; the total average earnings for the four incentive treatments were $17.1, $17.8, $27.1, and $26.9, respectively. The subjects were volunteer undergraduate students at the University of Haifa from all fields of study aged 23 years on average.

suggests, the authors argue, that experimenters ought to pay enough, or not pay at all. That is, while economists' belief in the effects of increasing financial incentives in experiments seems to be right on the money for reasonably high stakes, microscopic payments may induce behavior that, on the surface, contradicts their view of the world. Gneezy and Rustichini conjecture (and we agree) that their minimally paid NIS0.1 subjects might have been insulted by the microscopic compensation offered and as a result performed worse than the flat-fee subjects solving the tasks solely based on their intrinsic motivation.[2]

Importantly, Gneezy and Rustichini did not analyze their data at an individual level. As we show below, disaggregating the data allows for quantitative insights about the separate effects of incentives and abilities on cognitive performance. In Figure 1 we assume, as Gneezy and Rustichini did, that subjects were sampled from a common population, and plot, for each of the four incentive treatments, individual IQ performance against his/her IQ rank order. Hence an individual's IQ performance, measured on the vertical axis, induces his/her 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, connecting IQ-score observations yielded a "performance curve" for that treatment. While this illustrates the *within-treatment* variation in performance, one can similarly inspect the *across-treatment* variation by making comparisons among the performance curves. The following observations are noteworthy:

---

[2] Gneezy (2004) proposes that low-paid subjects may feel insulted or at odds with their self-perception. In addition, we note that the Gneezy and Rustichini experiment was not conducted double-blind, opening the door for reputation arguments a la Benabou and Tirole (2003).

First, notice that the performance curves for the high-incentive treatments (NIS1 and NIS3) are virtually identical and slope considerably upwards, implying that there is high within-treatment performance variation but hardly any across-treatment one. This result is most likely due to considerable within-treatment variation in cognitive abilities combined with hardly any across-treatment effort variation. One could also conceive that the large within-treatment performance variation is partly effort-driven, but large effort variation among subjects facing the same financial incentives seems unlikely. Thus ability (rather than incentive) differentials seem to determine performance differentials when incentives are high enough.

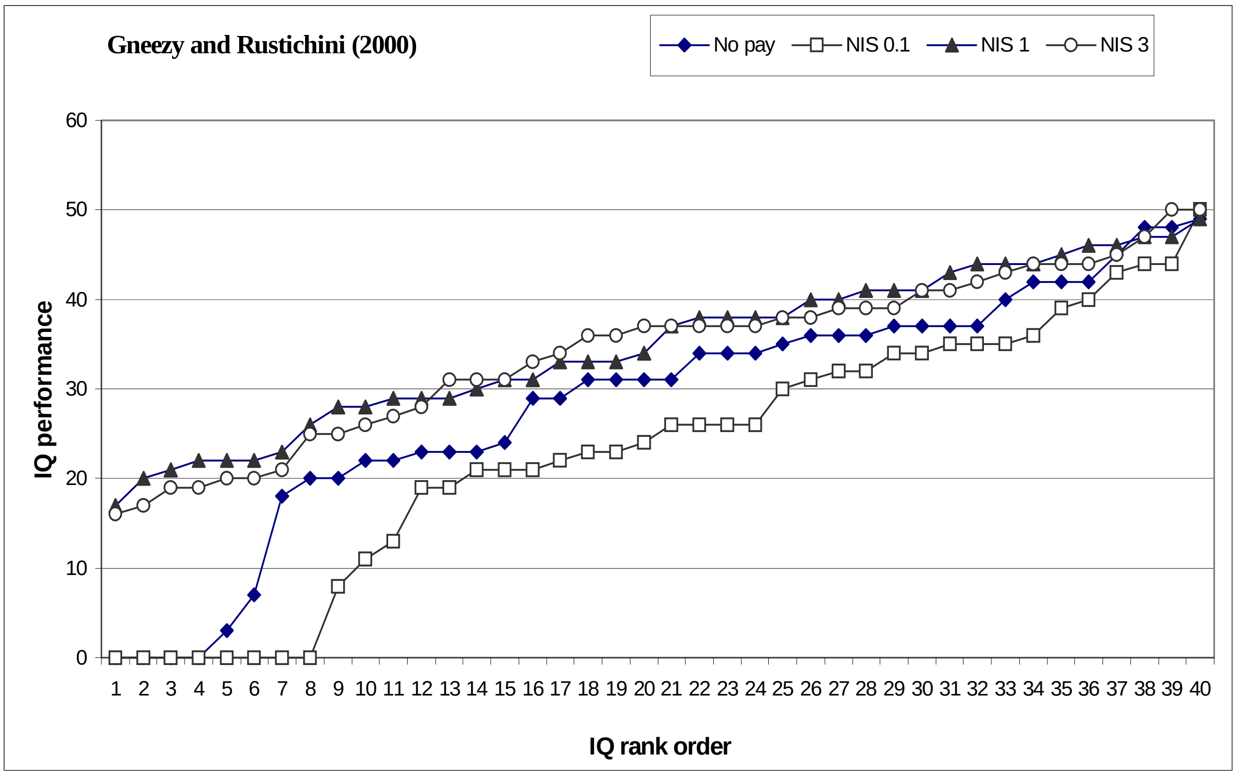

Figure 1: Individual IQ performance plotted against ascending IQ rank order separately for each of the four incentive treatments.

The last qualification leads us to examine the performance curves for the low-incentive treatments (no-pay and NIS0.1). The NIS0.1 performance curve lies persistently below the no-pay one. Further, the gap between them widens at the low-performance end, as does their distance from the two high-incentive performance curves. The first result is most likely due to general across-treatment effort (rather than ability) differentials. The second result, as correctly conjectured by Gneezy and Rustichini, suggests a specific kind of motivational problem at the low-performance end of the NIS0.1 treatment. The low-performance individuals are not necessarily the low-ability ones, but they seem to exhibit an outright refusal to perform, i.e. a different kind of "lack of motivation" that probably has little to do with "real" mental effort. Such reaction is much less prevalent in the no-pay treatment and entirely absent in the two high-incentive treatments.

Finally we combine the above insights. Leaving aside the motivational problems at the low-performance end, the performance variation is generally much greater within than across the four treatments. The largest across-treatment performance differential at the median IQ rank is 13 (24 vs. 37 correct answers in the NIS0.1 and NIS1 treatment, respectively). This is equivalent to the within-treatment performance differential between the first and third quartile of the NIS1 treatment (28 vs. 41), but note that the within-treatment performance differential is as high as 34 (16 vs. 50) for individuals ranked 1 and 40 in the two high-incentive treatments. Therefore, by assigning the (larger) within-treatment performance variation to ability differentials and the (smaller) across-treatment performance variation to effort differentials, we obtain our key result: ability differentials among individuals appear a much more powerful determinant of individual performance than do incentive differentials across treatments.

III. Discussion

As noted above, we do not know whether the unmotivated low-performance subjects in the no-pay and NIS0.1 treatments have low abilities; that is, these might be high-ability individuals. Yet such a scenario seems unlikely. First, high-ability subjects probably faced higher reputation costs which prevented them from dropping out (Benabou and Tirole, 2003). Second, one would need to explain who replaced the high-ability dropouts at the top-performance end: their lower-ability counterparts are unlikely candidates given their low financial stimulation and relatively high effort cost (in our view greater than for high-ability subjects). In any case, without the motivational problems (of possibly high-ability individuals), the dominance of ability over incentive effects would be even stronger.

One could conceive that Gneezy and Rustichini did not vary incentives sufficiently to assess the relative importance of ability and incentive effects. However, recalling the concurrence of the two high-incentive performance curves despite there being a big incentive differential (NIS1 vs. NIS3), further increasing the stakes seems unlikely to strengthen incentive effects. Further, the observation of Camerer and Hogarth (1999) that incentive effects can often be dampened by the existence of an artificial upper bound on performance does not seem to apply for the Gneezy and Rustichini data: only the very best subjects in each treatment reached the performance limit of 50 correct answers.

To conclude, we have offered a robust illustration of cognitive abilities being an important determinant of individual performance. Our findings have implications for experimental methodology. Among experimental 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. We add that experimental data should be interpreted conditional not only on the particular financial incentives employed, but also on the cognitive abilities of participants in the experiment.

To the extent that our illustration is representative of the relative importance of cognitive abilities, it also backs up Camerer and Hogarth's (1999) criticism of the labor-theoretic framework for its focus on the labor (effort) aspect and almost complete neglect of the capital (ability) aspect of cognitive production in experiments. In other words, we provide an empirical justification for integrating the insights of Camerer and Hogarth, as well as the motivational complications illustrated by Gneezy and Rustichini, into the labor theory of cognition (e.g., Smith and Walker, 1993). In a preliminary attempt at such a capital-labor framework, Wilcox (1993) proposes that subjects solve cognitive tasks by employing algorithms of various cognitive sophistication and effort cost. The next step is clearly to delve more deeply into what constitutes cognitive abilities and solution algorithms.

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.

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. 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.