Let´s be (intrinsically) honest: Introducing the revelation game

Author name: Stephen Mark Rosenbaum, PhD

Affiliation:
Institute for Marketing & Management
University of Southern Denmark
Campusvej 55, 5230 Odense M, Denmark

Tel: +45 6550 3262
Fax: +45 6615 5129
E-mail: sro@sam.sdu.dk

Abstract The role of norms of honesty in countervailing opportunistic behaviour in exchange interactions is almost universally recognized. Surprising though, while scholars have devised games for tapping into various other norms (such as cooperation, fairness, reciprocity, altruism and so forth), experiments tapping into norms of honesty, remain scarce. Those that do, suffer from various methodological limitations. In this paper, we introduce a novel economic experiment, the revelation game (RG), designed to tap into norms of honesty and circumvent these limitations. We report a RG experiment conducted in Uzbekistan, and discuss the implications of our findings.

Key words: Honesty, prosocial norms, experimental economics

1. Introduction/purpose

To paraphrase Arrow (1974), honesty toward strangers can be considered an important lubricant of any given society. However, while certain individuals exhibit honest behaviour by truthfully declaring tax liabilities (Wenzel, 2004) or handing in full wallets to police stations (West, 2010), others routinely deceive, lie, cheat and act dishonestly in a variety of social and economic interactions (Gneezy, 2005; Fischbacher and Heusi, 2008). In business settings, characterized by asymmetric information and incomplete contracts (Akerlof, 1970), an ability to distinguish honest types from dishonest ones, is crucial.

Why are some people honest, and others not? Two streams of thought dominate. First, the conventional economic model assumes that wealth-maximizing, self-regarding Homo Economicus is honest only insofar as the material rewards for honesty outweigh the incentives for being dishonest (Somanathan and Rubin, 2004). In the rational model of deterrence theory, individuals will thus be more inclined to be honest in situations when a) they stand to gain more from dishonest behaviour and/or, b) when the probability of being detected and the magnitude of punishment if caught, increase. Regulative changes (greater sentencing severity etc.) thus alter the formal incentives for honesty. Conversely, governmental corruptibility, ineffective contract and law enforcement raises payoffs to dishonest types relative to those of honest types (Somanathan and Rubin, 2004). This external cost-benefit view plays an important role in the theory of crime and punishment, which is the foundation for policies designed to prevent dishonesty (Mazar et al., 2008).

An alternative stream of literature attests to the presence of internalized norms of honesty which predominate purely selfish motives (Kahneman et al., 1986). Honesty is considered as a generalized value of acceptable behavior which is widely shared across a particular society (Pruckner and Sausgruber, 2009). Honesty norms are typically inculcated during childhood (Coles, 1997), such that people are socialized to be honest. There are also implicit costs and benefits to be honest in terms of being congratulated or ostracised by fellow group members (Somanathan and Rubin, 2004). Social norms thus create informal incentives for honesty. Honesty is thus part of an internal reward system such that people are willing to forgo financial benefits to obey norms (Mazar et al., 2008). In this internal view, policies designed to change normative behaviour can reduce levels of dishonesty in any given society. Economic experiments are typically used to examine the extent to which social norms countervail monetary incentives (Camerer, 2003). Experimental evidence robustly shows evidence of such other-regarding preferences, frequently toward complete strangers, despite incurring nontrivial monetary costs for prosocial behaviour (Kahneman et al., 1986; Ledyard, 1995; Camerer, 2003).

Surprisingly though, despite the importance of this endeavour and the vast array of other experiments tapping into norms such as cooperation, fairness, reciprocity, altruism, and so forth, previous attempts to tap into honesty norms have been somewhat scarce. The majority of experiments thus far have been field experiments, derivatives of the original "lost letter" experiments of psychologists (Milgram, 1963), including "lost dollar" (Yezer et al., 1996) and "lost wallet" (West, 2010) experiments which progressively increase the monetary value of the payoff for dishonest behaviour. Extensions within this genre include examining levels of honesty in terms of payments for newspapers in "honor stands" under different treatments of formal and informal norms (Pruckner and Sausgruber, 2009). However, there are a number of inherent drawbacks with field experiments including environmental contamination, the influence of peer monitoring, the risk of being exposed and so forth, which make it compelling to conduct lab experiments.

In this paper, we present a new and very simple experimental design that allows detecting dishonest behaviour where subjects face no threat of being exposed. We present an experiment where no formal incentives for honesty exist. Any abstention from lying, in terms of revealing the monetary discrepancy, must be interpreted as evidence of the existence of prosocial norms of honesty.

The remainder of the paper will be structured as follows. Section 2 will review the extant literature on honesty experiments. Section 3 introduces the experimental design and procedures of our novel Revelation Game conducted in Uzbekistan. Section 4 analyzes and discusses the results, while section 5 concludes.

2. Review of the experimental literature

Certain scholars have endeavoured to tap into honesty norms through a series of surveys detailing ethical dilemmas. For example, Frank et al. (1993) asked students in the US to imagine a situation where the owner of a small business is shipped ten computers, but is only invoiced for nine. Subjects were requested to indicate whether they would voluntarily reveal the billing discrepancy to the supplier if they were the owner. They found that 33% of all subjects reported the billing error to suppliers. However, there are inherent problems associated with such surveys – primarily the fact that there are no costs associated with misrepresenting type. In other words, individuals will naturally claim that they are always honest if dishonest behaviour could not be easily detected.

This led scholars to direct observation in field experiments. In his famous "lost letter" psychological experiments, Milgram (1963) found that seemingly lost letters were more likely to be returned to addressees such as "Medical Research Associates" than to "Friends of the Communist Party". However, the main disadvantage to psychological experiments is an inability to ascertain the extent to which individuals are motivated by monetary considerations. There are simply no monetary costs associated with being honest.

Subsequently, experimenters have conducted a number of "lost dollar" economic experiments, mainly in the field. In a derivative to Milgram´s (1963) lost letters, Yezer et al. (1996) leave 64 unsealed envelopes in graduate classes. The envelopes, fully-addressed and pre-stamped (yet with no indication of a return address), contain ten one-dollar bills together with a letter stating that the money was to pay back a loan (giving the impression that it had been written by a fellow student – invoking a sense of duty). They find that 44% of finders duly post the envelopes to the addressee, rather than pocketing their contents. Similarly, West (2005) dropped 20 wallets (each containing 20 USD) in mixed business and shopping districts in Tokyo and New York. The formal institutions (statutes) regarding lost and found situations are similar between the two countries (rules prescribing when the finder is obliged to hand in wallets, the fact that the finder is rewarded with the property if the item is not claimed within a similar specified period of time, and so forth). In Tokyo, 17 wallets were handed in, compared to New York, where 6 were handed in intact and 2 were handed in with their contents emptied. In another manifestation of this genre, Pruckner and Sausgruber (2009) report on a natural field experiment to examine the extent of honesty in the "honour system" of unmanned newspaper sales in Austria. They expect that people either pay full price or pay nothing at all. They conduct control ("The paper costs 60 cents") and two main treatments with internal- and external sanctions – appealing to honesty in a message next to the newspapers (The paper costs 60 cents. Thank you for being honest"), and reminding customers of the current legal norms ("The paper costs 60 cents. Stealing a paper is illegal"). They have two main findings. First, around two-thirds of customers for each category pay nothing, but the lowest level of pure free riding was witnessed in the moral treatment (63.4%) compared to control and legal treatment (67.5% and 66.7%, respectively). Second, there is a nontrivial number of people who pay the full amount, and this number is highest under the moral treatment (47%, N=118). Third, a sizable number of people pay something, but not the full price, and again, the amount paid is highest in the moral treatment (where the average non-zero payment is €0.38, compared to €0.16 in control and €0.15 in legal). They conclude that the effect of moral norms outweighs any external, legal incentive. Even if people do not contribute the full amount, they appear to exercise some form of guilty conscious compelling them to contribute something even in the absence of monitoring and external sanctions.

However, there are a number of problems associated with such field experiments. As Frank et al. (1996) acknowledge of their (1993) experiments detailed above, subjects may have simply ignored envelopes left in their classrooms, assuming that they contain nothing of value (Frank et al., 1996). Furthermore, subjects´ behaviour may have been influenced by the implied identity of the sender (a fellow student – which may induce more honesty), the presence of peer monitoring, and the absence of external incentives.

Bucciol and Piovesan (2011) examine honesty among children between the ages of 5-15 at a summer camp in Italy. In their control treatment, each child is asked to toss a fair coin once in private and asked to record the outcome on a paper sheet. The coin was white on one side and black on the other. The children were told that only those reporting white were rewarded (with ice cream and soda). The children thus had a deliberate incentive to cheat. Although the authors could not tell whether a child was honest or not, the observed outcome could be compared with the equal distribution implied by a fair coin. Not surprisingly, the authors find that white was reported in 85% of all cases (i.e. around 60% more than the 50% expectation),

yet less than 100% so. In their second treatment, children are explicitly requested to refrain from cheating. While this affected honesty levels among girls (16% fewer reported white), this was not the case for boys.

In an attempt to overcome some of the obstacles associated with field experiments (especially contamination of natural environments), attention has turned to the controlled conditions of the lab. In a variant of Bucciol and Piovesan´s (2011) field experiment, Fischbacher and Heusi (2008) report an experiment where subjects privately roll a 6-sided dice. They are informed that the number they report to have rolled is associated with a particular payoff – 1,2,3,4, and 5 CHF for the corresponding numbers, and zero if they shook a 6. The experimenter could not see what number they rolled, so the subjects could report whatever number they wanted. They then compared the answers to the underlying probability distribution of dice-rolls (i.e. 1/6 or 16.7% for each score). Found that 35% said that they had rolled a 5 and 27.2% who said that they had rolled a 4. Only 6.4% reported that they had rolled a 6 (zero payment). On the basis of this underlying probability distribution, they conclude that 39% of the subjects were honest, 22% lied completely, while the remainder lied, but did not maximize their income by doing so (typically by reporting a 4). They argue that an explanation for such behaviour in the absence of any external incentives is the desire to maintain a favourable self-concept, including honesty and non-greediness. This explanation relates to the theory of self-concept maintenance (Mazar et al., 2008). But, impossible to detect extent of lying with this design, compared to ours. There are however drawbacks to this study – the results are self-reported and compared to an underlying probability distribution (no direct observation of behaviour).

This study extends the existing literature by introducing the RG experiment, which, we propose, overcomes many of the limitations associated with previous experiments.

3. Introducing the Revelation Game

3.1. The game and experimental design

In essence, the revelation game taps into intrinsic honesty by identifying how many subjects voluntarily reveal a windfall payment to the experimenter. Very simply, the foreign researcher overpays each subject for their performance in previous experiments, stated clearly on an accompanying payout form, by a certain, standard amount of money.

Consistent with other economics experiments, the RG is conducted against the backdrop of neoclassical economics. The key dimensions of the RG set-up are designed to maximize the incentives to act dishonestly. In view of the experimental evidence from interactive or social settings suggesting that subjects base decisions on the costs not only to themselves, but also those imposed on the counterpart (Yezer et al., 1996; Gneezy, 2005), the RG features a strangers design employing double-blind anonymity with respect to both fellow subjects and the experimenter (Hoffman et al., 1994). In the RG, the counterpart is an apparently profligate foreign researcher, who has flown in, distributed large sums of money, and is due to fly out again immediately afterwards. The RG is a one-shot single decision-making situation so as not to arouse suspicion or to allow for an evolutionary process where a preference for honesty may emerge in repeated games. There are no opportunities for communication or peer monitoring, both of which mechanisms are documented to increase normative behaviour (Camerer, 2003). Moreover, the subject is led to believe that the payment is purely serendipitous, such that a non-revelation decision is not likely to result in a risk of detection. There are thus no external incentives to be honest. Moreover, there are relatively large monetary incentives associated with not revealing the discrepancy, representing nontrivial opportunity costs for behaviour which deviates from the sub-game perfect Nash equilibrium. On the basis of the above, the unique Nash equilibrium prediction for this game is for the subject not to reveal the windfall and simply pocket the cash.

3.2. Experimental procedure

The RG experiment was conducted with 40 students at the Westminster International University in Tashkent (WIUT) in Tashkent, Uzbekistan in February 2010. Subjects were recruited through both classes and posters, in which monetary rewards for participation were emphasized. No subject had ever participated in an experiment before, and only one subject had some prior knowledge of game theory. The demographic characteristics of the participating students are illustrated in Table 1.

Table 1: Demographic characteristics

Characteristics	Uzbekistan (N = 40)
Country of origin (total number)	Uzbek (35) Russian (3) Armenian (1) South Korean (1)
Mean age in years (standard deviation)	19.98 (1.42)
Number of men/women	29/11
Number of undergraduates/postgraduates	38/2
Main subjects of study	Economics/Business (33) Law (3) Business Computing (2) Others (2)

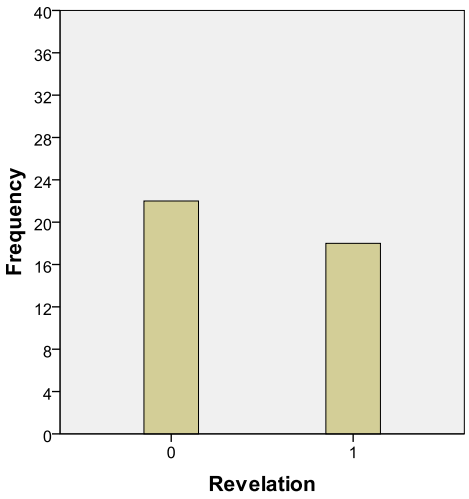
The revelation game is played as follows. The subjects had been randomly assigned an ID code, completed a pre-game questionnaire, played other games, and completed a post-game questionnaire. This marks the beginning of the RG experiment. Upon completion of the post-game questionnaire, subjects are accompanied individually to a private room to receive their remuneration from the foreign researcher. Only the foreign researcher and the subject were in the room. The remaining subjects are seated in a holding room under the supervision of a member of the experimental team and are not permitted to communicate with each other in any way. In the confines of the private payment room, each subject then receives a sealed envelope with corresponding ID number on. The envelope contains the payout form (which correctly advises them of their winnings from the previous games and show-up fee) and an amount in local currency (UZS). However, the monetary amount received exceeded the stated figure on the payout form by the equivalent of USD 5, constituting over half a day´s average wage in Uzbekistan (Asian Development Bank, 2009)[1]. The subjects were then provided with two instructions: (a) retire to a table at the other end of the room and count the money in private, and (b) if satisfied that the enclosed cash corresponded to the winnings declared on the payout form, to sign a declaration to this effect. Players were given time to count. As the student was facing away from the foreign researcher, the latter could unobtrusively observe that the counting process had begun, without the subject suspecting that he/she was being monitored. When the subject had signed the declaration form, he/she was then released from the payment room and escorted individually out of the building by an experimenter, before the next subject was admitted from the holding room. In the period between one student leaving the room and the next one arriving, the foreign researcher discretely marked an asterisk alongside the ID number to indicate an incidence of revelation, and the paper was then turned over to prevent any attempts by future entrants to link the presence or absence of an asterisk with variations in behaviour.

4. Results

The results from the RG experiment are depicted by the histograms in Figure 1. Fully 18 of the 40 Uzbek subjects (45%) revealed that they had received the windfall and returned their excess winnings to the experimenter, while 22 students simply pocketed the money.

[1] A key feature of the RG is that the subject is not aware that he or she is playing a game. It is important to note, however, that the RG does not involve deception, which is conventionally proscribed by experimental economists on the grounds that deceptive practices contaminate future subject pools (Hertwig and Ortmann, 2001). Consistent with the distinctions in the experimental literature (cf. McDaniel and Starmer, 1998; Hertwig and Ortmann, 2001), conveying false information to experimental subjects is deceptive, whereas withholding certain information from subjects is not (Hey, 1998). The revelation game is clearly an example of the latter.

Figure 1: RG behaviour in Uzbekistan (N=40)



Key: 0=non-revelation; 1=revelation

Given that homo economicus suffers no costs for dishonest behaviour, the study further corroborates previous experimental findings of behaviour that deviates significantly from Nash equilibrium predictions of pure self-interest. In terms of previous experiments, our findings are quite consistent with those of the lost dollar games where 44% on average exhibited honest behaviour (Yezer et al., 1996), while exceeding the 33% of subjects revealing invoicing discrepancies in the Frank et al. (1993) study. However, our results may be considered quite remarkable given the relative size of the stake (constituted over half a day´s national average wage) compared to a stake equivalent to two hours average salary in the Yezer et al. (1996) study and hypothetical money in the Frank et al. (1993) paper. Furthermore, the findings of the present study are quite remarkable in the light of the absence of overt monitoring (compared to, for example, the lost dollar experiments) and external sanctions. In the controlled conditions of the laboratory and the fact that subjects were provided with every incentive not to reveal the windfall, our results may suggest that there are greater numbers of individuals imbued with honesty norms than hitherto believed.

Subsequent Probit regressions to test the possible effect of demographic characteristics on the binary variable (revelation, nonrevelation) revealed that the effect of gender (0.573), age (0.599), semester (0.804), and ethnic background (0.784) were all highly insignificant. This suggests that the honesty evidenced in Tashkent is of an intrinsic nature, not correlated to any particular demographic trait.

5. Conclusion

In this paper, we introduce a novel experiment, the RG, to tap into predispositions to honesty. It differs from previous attempts to examine norms of honesty in a number of key dimensions. Compared to previous economic experiments based on self-reporting (cf. Fischbacher and Heusi, 2008), the RG allows for direct observations of behaviour. By providing subjects with optimal conditions to act in a self-interested manner and not reveal the windfall, we propose that the RG offers experimenters the opportunity to derive a greater insight into intrinsic honesty. Our findings that almost half of our experimental subjects in Uzbekistan revealed the serendipitous amount and duly returned it to the experimenter suggest high levels of honesty norms. This behaviour may be considered even more intrinsic given that we could not identify any demographic characteristic which had a significant effect on revelation behaviour.

As a novel experiment, the RG has certain advantages for experimenters. For example, it can easily be appended to other experimental sessions, as the payment process needs to take place under all circumstances. Furthermore, as an economic experiment, it is relatively cheap to run, as the monetary discrepancy has to be low enough so as not to arouse suspicion. Relatedly, it does not require elaborate instructions or written instructions, which run the risk of misunderstanding even when double back-translated.

Prior to closing, we would like to highlight some limitations with this study, which point to interesting areas of future research. First, the experiments were conducted in a single country, Uzbekistan. In this regard, there may be country-specific factors which caused the observed relatively high levels of honest behaviour, such as an inherent suspicion of covert monitoring in an authoritarian country. Future studies playing the RG in less-repressed societies may add significantly to our knowledge about honest behaviour in different societies. Second, our RG involved little social interaction between the foreign researcher and the experimental subjects. Future studies where the cashier was familiar to the subjects (for example, a teacher) would allow experimenters to examine the effects of stronger internal sanctions (such as guilt, embarrassment, as so forth) on honest behaviour. Similar extensions could include having two subjects in the room concurrently to examine the effects of peer monitoring on honesty.

Given the importance of distinguishing between honest and dishonest individuals in economic transactions, we believe that the revelation game offers an important contribution to the field. By overcoming the methodological limitations of previous studies, we propose that the RG is a step in the right direction of allowing experimenters to more accurately measure honest behaviour.

Acknowledgements

We are very grateful to Abdumalik Djumanov and his team for their invaluable assistance in running the experiments in Tashkent. Financial support from the University of Southern Denmark is gratefully acknowledged.

References

Akerlof, G.A. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84, 488-500.

Arrow, K. J. 1974. *The Limits of Organization*. Norton: New York.

Asian Development Bank., 2009. Key Indicators for Asia and the Pacific: Highlights. Manila, Philippines: Asian Development Bank.

Bucciol, A. and Piovesan, M. 2011. Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology* 32, 73-78.

Camerer, C.F., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Coles, R. 1997. *The Moral Intelligence of Children*. Oxford: Oxford University Press.

Fischbacher, U. and Heusi, F. 2008. Lies in Disguise: An experimental study on cheating. *Thurgau Institute of Economics* Research Paper 40.

Frank, R.H., Gilovich, T.D. and Regan, D.T. 1993. Does Studying Economics Inhibit Cooperation? *Journal of Economic Perspectives* 7, 159-171.

Frank, R.H., Gilovich, T.D. and Regan, D.T. 1996. Do Economists Make Bad Citizens? *Journal of Economic Perspectives* 10, 187-192.

Gneezy, U. 2005. The Role of Consequences. *American Economic Review* 95, 384-394.

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

Hey, J.D. 1998. Experimental economics and deception: A comment. *Journal of Economic Psychology* 19, 397-401.

Kahneman, D., Knetsch, J.L., Thaler, R., 1986. Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *American Economic Review* 76, 728-741.

Ledyard, J. 1995. Public goods: a survey of experimental research. In: J. Kagel and A.E. Roth (Eds.), *Handbook of Experimental Economics*. Princeton: Princeton University Press.

Mazar, N., Amir, O. and Ariely, D. 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research* XLV, 633-644.

McDaniel, T. and Starmer, C. 1998. Experimental economics and deception: A comment. *Journal of Economic Psychology* 19, 403-409.

Pruckner, G.J. and Sausgruber, R. 2009. Honesty on the Streets: A Natural Field Experiment on Newspaper Purchasing. *The Austrian Center for Labor Economics and the Analysis of the Welfare State*. Working Paper No. 0924.

Somanathan, E. and Rubin, P.H. 2004. The evolution of honesty. *Journal of Economic Behavior and Organization* 54, 1-17.
Wenzel, M., 2004. The Social Side of Sanctions: Personal and Social Norms as Moderators of Deterrence. *Law and Human Behavior* 28, 547-567.
West, M.D. 2010. *Law in everyday Japan: sex, sumo, suicide, and statutes*. Chicago, IL: University of Chicago Press.
Yezer, A.M., Goldfarb, R.S. and Poppen, P.J. (1996). Does Studying Economics Discourage Cooperation? Watch What We Do, Not What We Say or How We Play. *Journal of Economic Perspectives* 10, 177-186.