



## A field experiment on dishonesty: A registered replication of Azar et al. (2013)



Jakub Prochazka<sup>a,\*</sup>, Yulia Fedoseeva<sup>b</sup>, Petr Houdek<sup>c</sup>

<sup>a</sup> Masaryk University, Faculty of Economics and Administration, Department of Corporate Economy, Lipova 41a, 602 00 Brno, Czech Republic

<sup>b</sup> Masaryk University, Faculty of Economics and Administration, Department of Corporate Economy, Lipova 41a, 602 00 Brno, Czech Republic

<sup>c</sup> University of Economics Prague, Faculty of Business Administration, Department of Management, W. Churchill Sq. 1938/4, 130 67 Prague, Czech Republic

### ARTICLE INFO

#### Keywords:

Dishonesty  
field experiment  
pre-registered replication  
customer behaviour  
JEL classification: C93  
D12  
D91

### ABSTRACT

This study is a registered replication of a field experiment on dishonesty by Azar et al. (2013). Their main finding was that most customers of an Israeli restaurant did not return excessive change; however, customers who received a higher amount of excessive change returned it more often than people who received a lower amount. Our study, which was conducted on a sample of customers of restaurants in the Czech Republic ( $N=219$ ), replicated the results of the original study. The high excessive change condition increased the chance of returning the excess change by 21.7 percentage points (17.4 percentage points in the original study). The findings show that the psychological costs of dishonesty can outweigh its financial benefits. We similarly found that repeat customers and women were more likely to return the excessive change than one-time customers and men. The majority (70%) of customers in our sample returned the excessive change. We discuss the importance of field studies and replications of them in the further development of research into dishonest behavior.

### 1. Introduction

This study is a registered replication of a field experiment on dishonesty carried out in an Israeli restaurant by Azar et al. (Azar, Yosef and Bar-Eli, 2013). They found that most customers did not return excessive change and, contrary to the expected hypothesis, customers who received a higher amount of excessive change returned it more often than people who received a lower amount. This result calls into question the simple economic prediction that people will be more dishonest if they can make more money by cheating (Gneezy, 2005). On the contrary, the findings corroborate theories that people do not evaluate material benefits and costs only. They also care about their moral self-esteem, and there is a limited magnitude of dishonesty that a person is willing to bear (Barkan, Ayal and Ariely, 2015, Cohn, Maréchal, Tannenbaum and Zünd, 2019, Mazar, Amir and Ariely, 2008). Stealing a small sum may be easily overlooked, rationalized or excused; on the other hand, customers may not be able to maintain a positive self-image when stealing a more substantial amount of money. The original study also confirmed the suspected antecedents of dishonesty, e.g., repeat customers returned the excessive change much more often than one-time customers and women returned it more often than men, especially among repeat customers.

There are several reasons for the replication of the original field

experiment. First, the current literature on determinants of cheating and the magnitude of it is inconclusive. Current reviews demonstrate that people behave differently in the laboratory and in the field and in different tasks or situations that are used by researchers to elicit cheating behavior (Gerlach, Teodorescu and Hertwig, 2019, Jacobsen, Fosgaard and Pascual-Ezama, 2017, Rosenbaum, Billinger and Stieglitz, 2014). Findings from the laboratory may not be generalizable to real-world settings and/or results of field experiments may be too contextually idiosyncratic. Second, studies focusing on naturally occurring situations that allow for cheating are rare. Their results cannot be systematically analyzed, which prevents the formulation of any conclusions regarding the magnitude or determinants of dishonesty in the field. A case in point is the influence of cultural norms; some studies have found that individual dishonesty is not culture-dependent and not correlated with institutional honesty (Pascual-Ezama et al., 2015), others have found the exact opposite (e.g., Fisman and Miguel, 2007, Gächter and Schulz, 2016). Third, the replication crisis in the social sciences has showed that many of the published effects had been found by chance or due to the specific conditions of a study. A number of attempts to replicate seminal effects in the research on dishonesty have failed (Dimant, van Kleef and Shalvi, 2020, Kettle et al., 2017, Kristal et al., 2020, Schild, Heck, Ścigala and Zettler, 2019, Verschuere et al., 2018). Therefore, an effect found in a

\* Corresponding author

E-mail addresses: [jak.prochazka@mail.muni.cz](mailto:jak.prochazka@mail.muni.cz) (J. Prochazka), [438028@mail.muni.cz](mailto:438028@mail.muni.cz) (Y. Fedoseeva), [petr.houdek@vse.cz](mailto:petr.houdek@vse.cz) (P. Houdek).

single study should be viewed with caution.

Nevertheless, our replication, which was conducted in two restaurants in the Czech Republic, largely confirms the results of the original study. The high excessive change condition increased the chance of returning the extra change by 21.7 percentage points (17.4 percentage points in the original study), and we also found that repeat customers and women were more likely to return the excessive change than one-time customers and men. In our sample, most customers returned the excessive change (in the original study, 66% of customers were dishonest; in our study, 70% of customers were honest).

### 1.1. The need for replications of field experiments on dishonesty

In order to determine the external validity and robustness of laboratory or survey results, field experiments must be performed (Harrison and List, 2004, Zitzewitz, 2012). Self-report measures are based on participants' reflections on their behaviors across various real-life situations and naturally suffer from motivational biases. For example, Alem et al. (Alem, Eggert, Kocher and Ruhinduka, 2018) showed a substantial difference between stated ethical behavior in a survey and revealed behavior in a field experiment. Laboratory measures elicit participants' behavior in a novel, artificially structured environment and the findings may not be generalizable to a real-world situation (Levitt and List, 2007, Winking and Mizer, 2013). Field behavioral measures exploit participants' responses in naturally occurring situations; however, they could be full of contextually rich confoundings (Gneezy and Imas, 2017, Hauser, Linos and Rogers, 2017).

Although several studies have presented links between dishonest behavior in the laboratory and in the field, this association is only weak or moderate at best (Dai, Galeotti and Villeval, 2017, Franzen and Pointner, 2013, Potters and Stoop, 2016). Moreover, most of the experiments on dishonesty use samples of students, only a minority of studies are field experiments that have focused on dishonesty in everyday life (e.g. the lost wallet, the misdirected letter) or business transactions (e.g. fraudulent price discrimination, honour system payment, free riding on public transport). Field experiments are usually focused on the dishonesty of service providers in their interaction with customers (Balafoutas, Beck, Kerschbamer and Sutter, 2013, Busse, Israeli and Zettelmeyer, 2017, Conrads, Ebeling and Lotz, 2015, Dugar and Bhattacharya, 2016, Vranka et al., 2019). There is a lack of studies that focus on customers' dishonesty in their interaction with sellers (for exceptions, see (Buccioli, Landini and Piovesan, 2013, Potters and Stoop, 2016)). Azar et al. (Azar, Yosef and Bar-Eli, 2013) is one of the few field experiments that is interested in the dishonesty of customers in a common business transaction. Therefore, our replication contributes to this under-researched topic.

There are certain antecedents of dishonesty that are common for various studies such as monitoring or intrinsic cheating costs, and other antecedents have been identified only in some experiments (Gerlach, Teodorescu and Hertwig, 2019, Jacobsen, Fosgaard and Pascual-Ezama, 2017, Rosenbaum, Billinger and Stieglitz, 2014). As experimental studies have various procedures and different operationalizations of dishonesty, it is hard to determine whether the antecedent is context-determined (i.e. typical for a specific type of dishonesty, situation, sample, or culture) or has only been found by chance. Conceptual replications of laboratory experiments on dishonesty are quite common (see (Rosenbaum, Billinger and Stieglitz, 2014)). However, besides one study (Vranka et al., 2019) which used an improved procedure similar to Rabinowitz et al. (Rabinowitz et al., 1993) and Feldman (Feldman, 1968), we have not found any replication of a field experiment focused on dishonesty in an everyday business transaction. The replication of an original study is essential to provide further evidence about the effect on a different sample and in a different cultural environment.

### 1.2. The need for the replication of studies with surprising results

The experiment of Azar et al. (Azar, Yosef and Bar-Eli, 2013) was conducted on a homogenous sample (i.e. the guests of one Israeli restaurant) and has never been replicated. However, its conclusions are cited as being generally valid (see e.g., (Abeler, Becker and Falk, 2014, Irlenbusch and Villeval, 2015)). This confidence in the findings of a single experiment seems unwarranted as the replication crisis has showed that many of the published effects had been found by chance (Open Science Collaboration, 2015, Klein et al., 2018). In particular, attempts to replicate effects in dishonesty research have been unsuccessful on several occasions (Dimant, van Kleef and Shalvi, 2020, Kettle et al., 2017, Kristal et al., 2020, Schild, Heck, Ścigała and Zettler, 2019, Verschuere et al., 2018).

Lindsay (Lindsay, 2015) listed "a troubling trio" of factors leading to doubts about the reported effect and therefore increasing the need for further replication. One of these factors is that the study reported a surprising result. The effect of manipulation of the amount of excess change in the original study was surprising. It was in the opposite direction to that which the authors hypothesized and was not in line with previous research that showed either an opposite effect of stake size or temptation on dishonesty (Gneezy, 2005) or incoherent effects (Farrington and Kidd, 1977, Yuchtman-Yaar and Rahav, 1986).

Our main goal was to replicate the main effect of the experimental manipulation that was found in the original study and the test of the following hypothesis: Customers who receive more excessive change return it more often than customers who receive less excessive change.

In our replication study, we followed the original study and also observed some characteristics of customers (i.e., gender) and context variables (i.e. lunch/dinner) that had shown to be relevant predictors of the excessive change return rate in the original study. However, we did not state hypotheses about these variables and we only controlled for their effect to reduce possible confounding variables.

## 2. Method

This is a preregistered replication study. The registration is available online at: <https://aspredicted.org/blind.php?x=89e3br>.

### 2.1. Restaurants

We conducted the experiment in two restaurants in Brno, Czech Republic. Restaurant 1 is a large establishment close to the city centre specialized in burgers and traditional Czech cuisine. The restaurant can deliver around 50-70 meals at one time. Restaurant 2 specializes in Indian cuisine and can deliver around 20-30 meals at one time. Most main dishes in both restaurants cost around 150-250 CZK (Czech crowns approximately 6–10 Euros), which is similar to comparable middle-class restaurants in large Czech cities. On weekdays between 11 AM and 2 PM, discounted meals are available (for around 100 CZK per meal, which is approximately 4 Euros). Their typical customers are middle-class people between the ages of 20 and 60.

### 2.2. Procedure

During October 2019, waiters were asked to quasi-randomly return excessive change to customers who had paid their bill in cash. Waiters returned either an excess of 20 CZK (approximately 0.8 Euros) or an excess of 100 CZK (approximately 4 Euros). Waiters were asked to alternate the larger and smaller sums to customers who had paid cash (the original article does not state how the randomization was performed). In the original experiment waiters returned an excess of 10 Shekels (approximately 2.6 Euros) or 40 Shekels (approximately 10.5 Euros). For our replication study, we calculated the amounts taking into account the CZK/Shekel exchange rate and the difference in price level between similar Czech and Israeli restaurants. In line with the original

experiment, the aim was to ensure that the excess change was believable, rounded, and could easily be handled by the waiters. The waiters were given the instruction not to return excess change to participants who were clearly concentrated on other activities during the payment (e.g., making a phone call, in discussion with another person).

The research was conducted throughout the week during the operating hours of the restaurant, from 11 am until 10 pm. As is typical in some Czech restaurants, customers paid at the bar. One of the researchers was sitting near the bar at the table, secretly observing the payment and recording several variables: whether it was lunch or dinner, the gender of the waiter, the gender of the customer, whether the customer was clearly concentrating on other activities during the payment, i. e. whether the customer was checking the returned change.

After the customers had left the restaurant, the waiters recorded several variables: whether the excess change had been returned, if it was 20 or 100 CZK, whether the customer was visiting the restaurant for the first time or was a repeat customer (as far as they could remember), and the value of the bill. The variables that the experimenter observed and that the waiters reported were the same as in the original experiment of Azar et al. (Azar, Yosef and Bar-Eli, 2013).

### 2.3. Sample

We observed payments from 221 patrons of the restaurants. We excluded two customers who did not even look at the change. Therefore, the final sample consisted of 219 guests who had paid in cash and had received excessive change from the waiter. Of the sample, 120 (55%) received 20 CZK, and 99 received 100 CZK excess change. Following the preregistration, we stopped the data collection at the end of the day, shortly after we had reached 214 observations, as 214 observations allowed for testing 50% of the original effect size (i.e.  $d = .403$ ) with 90% test power on  $\alpha = .05$ .

### 2.4. Variables

We analyzed data on the seven variables shown in Table 1.

The variable *Return* is the dependent variable which indicates honest behaviour (the excessive change was returned) or dishonest behaviour (it was not returned). *High-change* is a manipulated experimental condition and refers to high (100 CZK) or low (20 CZK) excessive change that the guest received from the waiter. Other variables were observed by one of the researchers or a waiter. We observed and preregistered the same variables as the original study except for variables which were not significant predictors of dishonesty and at the same time a) which could not be measured in the conditions of our experiment or b) which would be complicated to record during busy working hours. We focused on the main effect of *High-change* and controlled for variables that had shown to be relevant in the original study (*Repeated*, *Woman-payer*) and on the variables that were easy to observe.

After the data collection had ended, we had to exclude from the analyses some variables that we had originally intended to observe due to their low variability. The first of these was the variable *Male-waiter*,

**Table 1**  
Description of variables.

Variable	Explanation
Return	1 if the excessive change was returned, 0 if not
High-change	1 if it was 100 CZK excess change, 0 if it was 20 CZK
Repeated	1 if it was a repeated customer of the restaurant, 0 otherwise
Weekend	1 if it was weekend, 0 otherwise
Evening	1 if the order was made between 17:00 and 21:59, 0 if it was between 11:00 and 16:59
Bill-per-person	Price in CZK per person
Woman-payer	1 if the bill was paid by a woman, 0 otherwise

**Table 2**  
Descriptive statistics of Return in various conditions.

	N	%	95%CI	
Low change	120	62%	53%	70%
High change	99	80%	72%	88%
Not repeated	143	66%	58%	74%
Repeated	76	78%	68%	87%
Workday	180	67%	60%	74%
Weekend	39	82%	69%	95%
Lunch	171	68%	61%	75%
Evening	48	75%	62%	88%
Man-payer	148	64%	56%	72%
Woman-payer	71	82%	72%	91%
Total	219	70%	64%	76%

Note. N = number of customer who returned excessive change in each condition; % = proportion of customer who returned excessive change in each condition; 95% CI = 95% confidence interval.

as a male waiter had handled the payment in only 3 of the 219 observations. The other excluded variables were multiple payers and bill sharing. In both restaurants, all the treated customers paid individually at the bar. Therefore, we did not observe multiple-payers or shared payments.

## 3. Results

### 3.1. Descriptive statistics

Table 2 shows frequencies for all binary variables and the descriptive statistics for *Return* for each condition. 153 customers out of 219 (70%) returned the excess change. The proportion of customers who returned the excess change was higher in the high-change condition (80%) than in the low-change condition (62%). This difference was significant,  $t(217) = 3.01$ ,  $p = .003$ , Fischer's exact test statistic = .0048,  $p < .01$ . The average *Bill-per-person* was 332.30 CZK ( $SD = 450.52$ ). Customers who did not return the excess change spent slightly more ( $M = 354.61$ ,  $SD = 630.94$ ) than customers who returned the excess change ( $M = 322.68$ ,  $SD = 347.21$ ),  $t(82.49) = .39$ ,  $p = .700$ .

### 3.2. Hypothesis testing

To test the hypothesis with control variables, we used the same linear regression model as Azar et al. (Azar, Yosef and Bar-Eli, 2013). We multiplied the dependent variable (i.e. *Return*) by 100, so the coefficients present the percentage increase in the probability of returning the excessive change. The model predicted *Return* of excess change significantly,  $F(6, 212) = 4.90$ ,  $p < .001$ , and explained 12.2% of the variance. Our analysis provided very similar results to the original study. The high-change condition increased the chance of returning excess change by 21.7 percentage points (17.4 percentage points in the original study), and this effect was significant. Therefore, we successfully replicated the effect from Azar et al. (Azar, Yosef and Bar-Eli, 2013) as we found support for the hypothesis that customers who receive more excessive change return it more often than customers who receive less excessive change.

### 3.3. Supplementary analyses

As in the original study, we found effects for *Repeated* and *Woman-payer* as repeated customers and women returned the excessive change more often than one-time customers and men. We also found weak and barely significant effects for *Weekend* and *Bill-per-person*, which were insignificant in the original study (see Table 3).

Azar et al. (Azar, Yosef and Bar-Eli, 2013) also included in their model the interaction between *High-change* and *Repeated*. They found

**Table 3**  
Linear probability model.

Variable	Full model			With moderation		
	B	β	P	B	β	p
Intercept	49.60		<.001	53.85		<.001
High-change	21.69	.24	<.001	13.81	.15	.064
Repeated customer	13.94	.14	.031	4.26	.04	.605
Weekend	16.55	.14	.047	16.74	.14	.043
Evening	13.67	.12	.095	14.39	.13	.078
Bill-per-person	-.02	-.16	.038	-.02	-.16	.034
Woman-payer	15.50	.16	.016	13.49	.14	.037
High-change * Repeated				23.57	.18	.065

Note. Model without moderation  $R^2 = .122$  ( $p < .001$ );  $\Delta R^2 = .014$  ( $p = .065$ ).

this interaction to be significant as the effect of *High-change* was stronger for repeated than for the one-time customer. We tested this effect in the second step of the regression model. As can be seen in Table 3, we found an effect in the same direction. In the low-change condition, one-time customers (61%) and repeated customers (63%) returned the excess change with a similar probability. In the high-change condition, only 71% of one-time customers but all repeated customers (100%) returned the money. This effect was significant only in a simple model with two predictors (i.e. *Returned* and *Repeated*,  $p = .038$ ) but not in the full model ( $p = .065$ ) due to the lower test power of the model with 7 predictors. Adding the interaction to the model suppressed the effect of *Repeated* (see Table 3) because the effect was marginal in the low-change condition.

It is important to note that linear regression is not the most appropriate analysis for binary outcomes and we used it to compare our results with the results of the original study. Therefore, we also present the results of logistic regression which showed the same effects and provided the same conclusions (see Table 4) as the linear regression.

#### 4. Discussion

We replicated the findings of the original study (Azar, Yosef and Bar-Eli, 2013). In the high excessive change condition, customers were more likely to return the money than in the low excessive change condition. The result is in line with the explanation that a more substantial financial reward for dishonesty is associated with higher psychological costs and that these psychological costs outweigh the financial benefits (Barkan, Ayal and Ariely, 2015, Cohn, Maréchal, Tannenbaum and Zünd, 2019, Mazar, Amir and Ariely, 2008). It could also be that with the higher excess change it is harder for customers to pretend that they did not return the money because of inattention. The possible gain from dishonest behavior might be outweighed by the higher perceived risk of revealing the dishonest behavior (Kajackaite and Gneezy, 2017).

Moreover, a higher amount of excessive change is also more visible, and there is a higher chance that a customer will notice it. This explanation is also supported by the fact that as the size of the bill increased, the probability that the customer returned the excessive change decreased. Nevertheless, we do not believe that inattention is

**Table 4**  
Logistic regression.

Variable	B	p	OR
Intercept	-.18	.512	.83
High-change	1.24	<.001	3.45
Repeated customer	.80	.025	2.22
Weekend	1.05	.042	2.87
Evening	.87	.069	2.39
Bill-per-person	.00	.024	1.00
Woman-payer	.92	.013	2.52

Note. Naglekerke  $R^2 = .182$ , OR odds ratios.

the decisive reason for the findings. Repeat customers returned the large excessive change in all the cases whereas one-time customers did not. There is no reason to expect repeat customers to be more attentive; on the contrary, they are more likely to trust the restaurant. However, they could be expected to have higher costs associated with cheating, because they would not want to return to the restaurant (or could not) after their cheating had been detected. Similarly, they are likely to have a more positive attitude towards the restaurant and, therefore, dishonesty would result in higher psychological costs. Nevertheless, it can be assumed that regular customers would also have a better idea of how much the meal cost and might be expected to have the exact amount of money ready to pay. The higher excess change would be more noticeable.

In contrast to the original findings, in our sample, most customers returned the excessive change. Several factors may have caused this difference. In our experiment, customers paid at the bar, while in the original study, they paid at their table. Since the customer was alone with the waiter at the bar, we might expect the opposite effect due to fewer watching eyes. People standing at the bar also were required to make their decision immediately, whereas in the original study they had time to decide whether to return the excessive change while they remained seated at their table. According to the literature on default honesty, customers paying at a bar should be more honest (Capraro, Schulz and Rand, 2019). Similarly, at the table, the paying customer may be being observed by other people, and therefore, the cost of that person's dishonesty is higher. On the other hand, the presence of other people may help to justify the dishonest behavior through the diffusion of responsibility. Furthermore, personal proximity in the case of payments at the bar could elicit affinity and moral emotions (Greene et al., 2004) or the fear of revealing dishonesty (Kajackaite and Gneezy, 2017), both of which support honest behavior.

Another factor is cultural differences (Pascual-Ezama et al., 2015, Gächter and Schulz, 2016). The Czech Republic is a post-communist country and it now ranks 44th in the Corruption Perceptions Index (the higher the country ranks, the lower are the perceived corruption rates) whereas Israel ranks 35th (Transparency International 2019). Nevertheless, a laboratory study on a sample of Czech students revealed quite high preferences for honesty (Novakova, Houdek, Jolič and Flegr, 2018). In any case, the number of field experiments is insufficient for any exact cross-country comparisons. Dishonesty research generally suffers from similar problems as most experimental economics: the studies are mostly conducted in 'WEIRD' countries (Henrich, Heine and Norenzayan, 2010) and they are subject to severe publication bias (Gerlach, Teodorescu and Hertwig, 2019).

We also found that women are more honest than men. These results fit into the systematic evidence that women behave slightly more honestly than men in most situations (Gerlach, Teodorescu and Hertwig, 2019, Jacobsen, Fosgaard and Pascual-Ezama, 2017). Furthermore, customers were more honest on a weekend. We might speculate that customers are in a better mood at the weekend and are therefore less inclined to steal. There are also fewer people in restaurants at the weekend and the customers might expect that the waiter would pay more attention to their payment.

##### 4.1. Limitations

Our study was a close replication of the experiment of Azar et al. (Azar, Yosef and Bar-Eli, 2013). There were several differences between the two studies that originated from the distinctions between Czech and Israeli restaurants. Our aim was to replicate the main manipulation with excessive change in similar conditions, but in a different cultural context, controlling for the effect of variables that showed to be significant in the original study. The differences in payment (i.e., at the table or the bar), the slightly larger differences between the high and low excessive change conditions and the reduced number of control variables in the analyses should be taken into consideration when

comparing the results.

It is not clear from the original study how the authors randomized the experimental condition. We were not able to enforce randomization, so we asked waiters to at least alternate the larger and smaller sums of excess change to customers who had paid cash. However, the waiters were not able to maintain this alternation regularly, especially during the busy hours, and returned the cash quasi-randomly. We believe that this did not seriously affect the results, as the waiters did not know the hypotheses and had no interest in the outcome of the study. However, a possible unconscious bias on the part of waiters, could have affected the results.

We conducted our experiment in two restaurants in one Czech city. Our sample was not large and representative enough to enable a broad generalization of our results. However, we found a similar effect to the original study, which indicates that the effect of high/low excess change might exist across cultures. Nevertheless, further replications in various contexts are needed to support this conclusion.

## Conclusion

Our replication, which was performed on a different sample in another country, yielded the same results as the original study. We successfully replicated the effect from Azar et al. (Azar, Yosef and Bar-Eli, 2013) as we found support for the hypothesis that customers who receive more excessive change return it more often than customers who receive less excessive change. However, the results may be explained by several alternative explanations; future research should focus on differentiating their effects.

We argue that direct and conceptual replications are needed to understand the generalizability of the effects and potential antecedents of dishonest behavior (Pierce and Balasubramanian, 2015). Without field experiments and replications of them, there are few ways to build proper knowledge about factors that influence dishonesty in real-life situations.

## Funding

This work was supported by the Czech Grant Agency [grant number 18-13766S] and Masaryk University [grant numbers MUNI/A/1148/2018, MUNI/A/1073/2019].

## Declarations of Competing Interest

No

## Acknowledgements

Yulia Fedoseeva participated in the research as a part of her diploma thesis.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.socec.2020.101617](https://doi.org/10.1016/j.socec.2020.101617).

## References

- Abeler, J., Becker, A., Falk, A., 2014. Representative evidence on lying costs. *Journal of Public Economics* 113, 96–104. <https://doi.org/10.1016/j.jpube.2014.01.005>.
- Alem, Y., Eggert, H., Kocher, M.G., Ruhinduka, R.D., 2018. Why (field) experiments on unethical behavior are important: Comparing stated and revealed behavior. *Journal of Economic Behavior & Organization* 156, 71–85. <https://doi.org/10.1016/j.jebo.2018.08.026>.
- Azar, O.H., Yosef, S., Bar-Eli, M., 2013. Do customers return excessive change in a restaurant? A field experiment on dishonesty. *Journal of Economic Behavior & Organization* 93 (0), 219–226. <https://doi.org/10.1016/j.jebo.2013.03.031>.
- Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2013. What drives taxi drivers? A field experiment on fraud in a market for credence goods. *The Review of Economic*

- Studies* 80 (3), 876–891. <https://doi.org/10.1093/restud/rds049>.
- Barkan, R., Ayal, S., Ariely, D., 2015. Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology* 6, 157–161. <https://doi.org/10.1016/j.copsy.2015.08.001>.
- Buccioli, A., Landini, F., Piovesan, M., 2013. Unethical behavior in the field: Demographic characteristics and beliefs of the cheater. *Journal of Economic Behavior & Organization* 93, 248–257. <https://doi.org/10.1016/j.jebo.2013.03.018>.
- Busse, M.R., Israeli, A., Zettelmeyer, F., 2017. Repairing the damage: The effect of price knowledge and gender on auto repair price quotes. *Journal of Marketing Research* 54 (1), 75–95. <https://doi.org/10.1509/jmr.13.0291>.
- Capraro, V., Schulz, J., Rand, D.G., 2019. Time pressure and honesty in a deception game. *Journal of Behavioral and Experimental Economics* 79, 93–99. <https://doi.org/10.1016/j.socec.2019.01.007>.
- Cohn, A., Maréchal, M.A., Tannenbaum, D., Zünd, C.L., 2019. Civic honesty around the globe. *Science* 365 (6448), 70–73. <https://doi.org/10.1126/science.aau8712>.
- Conrads, J., Ebeling, F., Lotz, S., 2015. (Dis-)honesty: Measuring overcharging in a real-world market. *Journal of Behavioral and Experimental Economics* 57, 98–102. <https://doi.org/10.1016/j.socec.2015.05.003>.
- Dai, Z., Galeotti, F., Villeval, M.C., 2017. Cheating in the lab predicts fraud in the field: An experiment in public transportation. *Management Science*. <https://doi.org/10.1287/mnsc.2016.2616>.
- Dimant, E., van Kleef, G.A., Shalvi, S., 2020. Requiem for a Nudge: Framing effects in nudging honesty. *Journal of Economic Behavior & Organization* 172, 247–266. <https://doi.org/10.1016/j.jebo.2020.02.015>.
- Dugar, S., & Bhattacharya, H. (2016). Fishy behavior: Evaluating preferences for honesty in the marketplace. Available at SSRN 2726404. doi: 10.2139/ssrn.2726404.
- Farrington, D.P., Kidd, R.F., 1977. Is financial dishonesty a rational decision. *British Journal of Social and Clinical Psychology* 16 (2), 139–146. <https://doi.org/10.1111/j.2044-8260.1977.tb00209.x>.
- Feldman, R.E., 1968. Response to compatriot and foreigner who seek assistance. *Journal of Personality and Social Psychology* 10 (3), 202–214. <https://doi.org/10.1037/h0026567>.
- Fisman, R., Miguel, E., 2007. Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy* 115 (6), 1020–1048. <https://doi.org/10.1086/527495>.
- Franzen, A., Pointner, S., 2013. The external validity of giving in the dictator game. *Experimental Economics* 16 (2), 155–169. <https://doi.org/10.1007/s10683-012-9337-5>.
- Gächter, S., Schulz, J.F., 2016. Intrinsic honesty and the prevalence of rule violations across societies. *Nature* 531 (7595), 496–499. <https://doi.org/10.1038/nature17160>.
- Gerlach, P., Teodorescu, K., Hertwig, R., 2019. The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin* 145 (1), 1–44. <https://doi.org/10.1037/bul0000174>.
- Gneezy, U., 2005. Deception: The role of consequences. *American Economic Review* 95 (1), 384–394. <https://doi.org/10.1257/0002828053828662>.
- Gneezy, U., Imas, A., 2017. Lab in the field: Measuring preferences in the wild (Chapter 10). In: Banerjee, A.V., Duflo, E. (Eds.), *Handbook of Economic Field Experiments* (Vol. 1, pp. 439–464): North-Holland, Eds. .
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D., 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44 (2), 389–400. <https://doi.org/10.1016/j.neuron.2004.09.027>.
- Harrison, G.W., List, J.A., 2004. Field experiments. *Journal of Economic Literature* 42 (4), 1009–1055. <https://doi.org/10.1257/0022051043004577>.
- Hauser, O.P., Linos, E., Rogers, T., 2017. Innovation with field experiments: Studying organizational behaviors in actual organizations. *Research in Organizational Behavior*. <https://doi.org/10.1016/j.riob.2017.10.004>.
- Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world. *Behavioral and Brain Sciences* 33 (2-3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>.
- Irlenbusch, B., Villeval, M.C., 2015. Behavioral ethics: how psychology influenced economics and how economics might inform psychology. *Current Opinion in Psychology* 6, 87–92. <https://doi.org/10.1016/j.copsy.2015.04.004>.
- Jacobsen, C., Fosgaard, T.R., Pascual-Ezama, D., 2017. Why do we lie? A Practical guide to the dishonesty literature. *Journal of Economic Surveys*. <https://doi.org/10.1111/joes.12204>.
- Kajackaite, A., Gneezy, U., 2017. Incentives and cheating. *Games and Economic Behavior* 102, 433–444. <https://doi.org/10.1016/j.geb.2017.01.015>.
- Kettle, S., Hernandez, M., Sanders, M., Hauser, O., Ruda, S., 2017. Failure to CAPTCHA attention: Null results from an honesty priming experiment in Guatemala. *Behavioral Sciences* 7 (2), 28. <https://doi.org/10.3390/bs7020028>.
- Klein, R.A., Vianello, M., Hasselman, F., Adams, B.G., Adams, R.B., Alper, S., Nosek, B.A., 2018. Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science* 1 (4), 443–490. <https://doi.org/10.1177/2515245918810225>.
- Kristal, A.S., Whillans, A.V., Bazerman, M.H., Gino, F., Shu, L.L., Mazar, N., Ariely, D., 2020. Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences*, 201911695. <https://doi.org/10.1073/pnas.1911695117>.
- Levitt, S.D., List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21 (2), 153–174. <https://doi.org/10.1257/jep.21.2.153>.
- Lindsay, D.S., 2015. Replication in Psychological Science. *Psychological Science* 26 (12), 1827–1832. <https://doi.org/10.1177/0956797615616374>.
- Mazar, N., Amir, O., Ariely, D., 2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* 45 (6), 633–644. <https://doi.org/10.1509/jmkr.45.6.633>.

- Novakova, J., Houdek, P., Jolič, J., Flegr, J., 2018. Experimental assessment of cheating prevalence in the Czech Republic as an example of a post-communist country. *Deviant Behavior* 39 (9), 1217–1228. <https://doi.org/10.1080/01639625.2017.1410381>.
- Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. *Science* 349 (6251). <https://doi.org/10.1126/science.aac4716>. aac4716.
- Pascual-Ezama, D., Fosgaard, T.R., Cardenas, J.C., Kujal, P., Veszteg, R., Gil-Gómez de Liaño, B., Brañas-Garza, P., 2015. Context-dependent cheating: Experimental evidence from 16 countries. *Journal of Economic Behavior & Organization*, 116 379–386. <https://doi.org/10.1016/j.jebo.2015.04.020>.
- Pierce, L., Balasubramanian, P., 2015. Behavioral field evidence on psychological and social factors in dishonesty and misconduct. *Current Opinion in Psychology*, 6 70–76. <https://doi.org/10.1016/j.copsyc.2015.04.002>.
- Potters, J., Stoop, J., 2016. Do cheaters in the lab also cheat in the field. *European Economic Review* 87, 26–33. <https://doi.org/10.1016/j.euroecorev.2016.03.004>.
- Rabinowitz, F.E., Colmar, C., Elgie, D., Hale, D., Niss, S., Sharp, B., Sinclitico, J., 1993. Dishonesty, indifference, or carelessness in souvenir shop transactions. *The Journal of Social Psychology* 133 (1), 73–79. <https://doi.org/10.1080/00224545.1993.9712120>.
- Rosenbaum, S.M., Billinger, S., Stieglitz, N., 2014. Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45 181–196. <https://doi.org/10.1016/j.joep.2014.10.002>.
- Schild, C., Heck, D.W., Ścigała, K.A., Zettler, I., 2019. Revisiting REVISE: (Re)Testing unique and combined effects of REminding, Visibility, and SELF-engagement manipulations on cheating behavior. *Journal of Economic Psychology*. <https://doi.org/10.1016/j.joep.2019.04.001>.
- Transparency International (2019). *Corruption Perceptions Index 2019*. Retrieved April 30, 2020 from <https://www.transparency.org/cpi2019>.
- Verschuere, B., Meijer, E.H., Jim, A., Hoogesteyn, K., Orthey, R., McCarthy, R.J., Yıldız, E., 2018. Registered replication report on Mazar, Amir, and Ariely (2008). *Advances in Methods and Practices in Psychological Science* 1 (3), 299–317. <https://doi.org/10.1177/2515245918781032>.
- Vranka, M., Frollová, N., Pour, M., Novakova, J., Houdek, P., 2019. Cheating customers in grocery stores: A field study on dishonesty. *Journal of Behavioral and Experimental Economics* 83, 101484. <https://doi.org/10.1016/j.socec.2019.101484>.
- Winking, J., Mizer, N., 2013. Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior* 34 (4), 288–293. <https://doi.org/10.1016/j.evolhumbehav.2013.04.002>.
- Yuchtman-Yaar, E., Rahav, G., 1986. Resisting small temptations in everyday transactions. *The Journal of Social Psychology* 126 (1), 23–30. <https://doi.org/10.1080/00224545.1986.9713565>.
- Zitzewitz, E., 2012. Forensic economics. *Journal of Economic Literature* 50 (3), 731–769. <https://doi.org/10.1257/jel.50.3.731>.