University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 280

Honesty in the Digital Age

Alain Cohn, Tobias Gesche and Michel Maréchal

February 2018

Honesty in the Digital Age*

February 2018

Alain Cohn, Tobias Gesche and Michel Maréchal

Abstract

Modern communication technologies enable efficient exchange of information, but often sacrifice direct human interaction inherent in more traditional forms of communication. This raises the question of whether the lack of personal interaction induces individuals to exploit informational asymmetries. We conducted two experiments with 866 subjects to examine how human versus machine interaction influences cheating for financial gain. We find that individuals cheat significantly more when they interact with a machine rather than a person, regardless of whether the machine is equipped with human features. When interacting with a human, individuals are particularly reluctant to report unlikely favorable outcomes, which is consistent with social image concerns. The second experiment shows that dishonest individuals prefer to interact with a machine when facing an opportunity to cheat. Our results suggest that human interaction is key to mitigating dishonest behavior and that self-selection into communication channels can be used to screen for dishonest people.

Keywords: Cheating, honesty, private information, communication, digitization, lying costs

JEL Classification: C99, D82, D83

*Cohn: adcohn@umich.edu, School of Information, University of Michigan
Gesche: tgesche@fas.harvard.edu.com, Department of Economics, Harvard University & University College London
Maréchal: michel.marechal@econ.uzh.ch, Department of Economics, University of Zurich

We thank Gioa Birukoff, Matthias Fehlmann, Nicole Honegger, Ezra In-Albon, Elia In-Albon, Wolfram Ritter and in particular Cornelia Schnyder for excellent research assistance. We acknowledge the support by the Gottlieb Duttweiler Institute. Tobias Gesche also acknowledges support by the Swiss National Science Foundation (SNSF Grant #P2ZHP1_171900).

Introduction

Technological progress has radically transformed the way we communicate and interact with each other. For example, employees increasingly collaborate from remote places without physically meeting each other (e.g., Mateyka et al., 2012; Bureau of Labor Statistics, 2016), and retailers are closing their brick and mortar stores to sell their products online (e.g., Hortaçsu and Syverson, 2015; U.S. Department of Commerce, 2017). While modern communication technologies enable us to connect more easily with each other over great distances, they have also largely displaced face-to-face interactions and thereby reduced the "human touch" in our social and economic relationships (e.g., Turkle, 2012). In fact, recent developments in artificial intelligence (e.g. chatbots) may even completely replace human interaction in industries that have traditionally placed a strong emphasis on building close relationships, such as banking and insurance (Brewster, 2016; Hall, 2017). However, informational asymmetries between interacting parties characterize many of these relationships, creating opportunities for manipulation and deception. This raises the question of how human-machine interaction affects people's tendency to lie and cheat.

Why should individuals be more or less likely to cheat when interacting with a machine rather than a human being? Research in economics and other social sciences suggests that people are often motivated by social image, i.e., they care about how others perceive them (see Bursztyn and Jensen, 2017, for a recent overview). In the absence of a human counterpart, individuals may feel less observed and therefore may pay less attention to their image. Indeed, some studies indicate that people are more likely to adhere to social norms, i.e., to behave in ways that are deemed socially appropriate, when an audience is present (e.g., Kurzban et al., 2007; Andreoni and Bernheim, 2009). Thus, if individuals strive to be perceived as honest, they should feel less comfortable telling a lie when interacting with a person instead of a machine.

What if the machine is made more human, i.e., it is capable of mimicking (at least to some degree) a real person? According to social presence theory from social psychology (Short et al., 1976), people are sensitive to the perceived presence or "realness" of the interaction partner, which depends on the number of human cues (e.g., vocal pitch and gestures) that are transmitted through a communication medium.[1] Thus, this theory makes a more nuanced prediction about the role of social image in human-machine

[1] The label "social presence" is unfortunate in this context because the theory is about the perceived realness of the interaction partner rather than the actual presence of another person.

interactions. Specifically, a machine that is equipped with human features may activate individuals' image concerns and thereby reduce their propensity to cheat.

We test these two hypotheses in a controlled online experiment in which subjects could increase their earnings up to 20 Swiss francs (about US $20) by cheating on a coin-tossing task. Specifically, we asked subjects to flip a coin ten times, report their outcomes to the experimenter, and paid them according to their alleged success rate. Subjects performed the coin tosses from a remote place (typically from home) and reported their outcomes via the communication software Skype. We varied two factors using a two-by-two factorial design: (i) whether subjects reported their outcomes to a person or a machine, and (ii) whether the interaction involved oral or written communication. In treatment CALL, subjects had to call the experimenter on Skype to report their outcomes. We instructed them to make the calls without video to preserve the same degree of anonymity across conditions and thus ensure a ceteris paribus comparison of the treatments. In treatment FORM, subjects had to type their outcomes into a non-interactive online form. Thus, while subjects in treatment CALL interacted with a person, there was no human counterpart in treatment FORM. Treatment ROBOT was identical to treatment CALL, with the exception that subjects communicated with an interactive automated voice response system that prompted them to report their outcomes using the experimenter's pre-recorded voice messages. Thus, relative to treatment CALL, we only varied the presence of a real person while keeping the communication mode constant across conditions. Finally, we also conducted treatment CHAT where subjects had to report their outcomes to the experimenter using Skype instant messaging. The comparison of treatments CALL and CHAT allows us to test whether the transmission of richer human cues (i.e., voice instead of just text) affects people's tendency to cheat in human-human interactions.

Some design features are worth noting before we turn to the results. First, we designed the experiment so that the reporting was semantically identical across the four conditions, permitting a ceteris paribus comparison of the treatments. We also informed the subjects that they would not be asked any questions about the number of successful coin flips they report. Thus, subjects did not have to worry about justifying their claims, even when they interacted with a person. Second, because subjects performed the coin-tossing task in private, there was no way the experimenter (nor anyone else) could unambiguously identify whether a specific subject cheated. There was thus no actual risk of getting caught for misreporting unsuccessful coin flips. Finally, we gave subjects enough time in all conditions to contemplate on the number of successful coin flips they wished to report before they made their decisions. Thus, any differences between the human and machine conditions cannot be explained by greater time pressure in the human conditions, a factor that has been previously shown to affect people's tendency to cheat (e.g., Shalvi et al., 2012; Capraro, 2017; Lohse et al., 2018)

We find that subjects reported 54% successful coin flips, on average, in treatment CALL. This corresponds to a cheating rate of 8% if we assume that none of the subjects misreported their successful coin flips. By contrast, those in treatment FORM reported 62% successful coin flips, on average, which corresponds to a cheating rate of 24%. Thus, the cheating rate is about three times greater in FORM than in CALL. The two treatments not only differ in the presence of a human counterpart, but also in terms of media richness (voice vs. text). Treatment ROBOT enables a clean identification of the role of human presence in honest behavior. In this treatment, subjects reported 60% successful coin flips, on average, meaning that they cheated to a similar degree as in treatment FORM. Thus, making a machine more human by adding human voice does not render people more honest. Finally, in treatment CHAT, subjects reported 56% successful coin flips, on average, and were therefore similarly honest as those in treatment CALL. This strengthens the evidence that interaction with a real person rather than just human cues is key to promoting honest behavior.

Further analysis supports the notion that individuals' image concerns induce dishonest behavior in human-machine interaction. In particular, we find that subjects were almost three times as likely to report a high, and therefore suspicious, success rate (i.e., 8, 9, or 10 successful coin tosses) when reporting to a machine rather than to a human being. By contrast, more plausible success rates (i.e., 6 or 7 successful coin tosses) were reported with similar frequency across human and machine conditions. Overall, these findings are consistent with individuals who behave honestly because they care about other people's impression of them.

The results of the first experiment raise the question of whether dishonest people anticipate feeling less comfortable misrepresenting information when interacting with a person rather than with a machine. In other words, is it possible to screen for dishonest people by offering different communication channels that vary by whether or not a real person is at the other end of the line? To find out, we conducted a second experiment with new subjects who were given the choice between the online form and calling the experimenter on Skype to report the outcomes of their coin flips. Before making that choice, we elicited their propensity to cheat using the same coin tossing task. Thus, subjects performed the same task twice within one week – once under identical conditions to provide a proxy of their tendency to cheat and the second time to elicit their preferred reporting channel.

When asked to choose between communication channels, we find that subjects were about equally likely to select the call and the online form (50.6% vs. 49.4%). This suggests that the two reporting methods were, on average, perceived as similarly convenient. However, alleged cheaters, i.e., those who reported a high success rate in the initial coin tossing task, were significantly more likely to choose the online form for the second coin tossing task. Our estimates suggest that a person who claims the maximum payoff (i.e., winning on every single coin flip) is 19 percentage points more likely to select the online form compared to someone who reports 50% successful coin flips (the most likely outcome). Thus, more dishonest individuals avoid human interaction when there is an opportunity to cheat. This "selection on moral hazard" raises the possibility for firms and government agencies to screen for dishonest people. For example, they could improve their auditing procedures by offering a choice between different communication channels and targeting the suspicious cases of individuals who chose not to interact with a representative.

Our paper relates to several strands of the literature. First, our findings contribute to a growing literature in economics arguing that people strive to be perceived positively by others, even for non-instrumental reasons, and that these social image concerns can affect a wide range of behaviors, including charitable giving (e.g., Ariely et al., 2009; DellaVigna et al., 2012), labor supply (e.g., Kosfeld and Neckermann, 2011), voting behavior (e.g., DellaVigna et al., 2017), and consumption choices (e.g., Bursztyn et al., 2017). Social image concerns have also recently been incorporated into theoretical models of honesty to explain why many people do not exploit cheating opportunities to the full extent, even if they cannot get caught (Abeler et al., 2016; Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2016; Khalmetski and Sliwka, 2017). Our results suggest that people indeed like to be perceived as honest, and that the mere presence of a stranger at the other end of the line induces them to behave more honestly. In this sense, our paper also adds to the rapidly growing literature on the social and psychological motives of honest behavior (e.g., Gneezy, 2005; Mazar et al., 2008; Irlenbusch and Villeval, 2015; Shalvi et al., 2015; Abeler et al., 2016; Gächter and Schulz, 2016; Cohn and Maréchal, 2016).

Second, there is a literature on the evolutionary origins of pro-sociality arguing that our ancestors' living circumstances shaped human psychology in a lasting manner that now induces us to behave altruistically and honestly even towards genetically unrelated strangers (e.g., Dawkins, 2006; Trivers, 2006). The idea is that humans evolved in small groups where repeated interactions were common and people therefore had strong reputational incentives to behave pro-socially. These reputational concerns became so deeply ingrained that even the slightest cues of being observed by others can trigger pro-social behavior. Indeed, several studies show that even subtle human cues (e.g., an image of watching eyes) increase people's propensity to act altruistically and honestly (e.g., Haley and Fessler, 2005; Bateson et al., 2006; Ernest-Jones et al., 2011). However, the evolutionary legacy hypothesis has also been contested as other studies failed to replicate the original findings (e.g., Fehr and Schneider, 2010; Cai et al., 2015; Northover et al., 2017). Our results from treatment ROBOT suggest that vocal cues are not sufficient to activate people's reputational or image concerns. Instead, they point to the importance of real-time human interaction.

Finally, our paper also connects to a long-standing literature studying the impact of communication on economic behavior, such as coordination (e.g., Cooper et al., 1992; Crawford, 1998), cooperation (e.g., Isaac and Walker, 1988; Bicchieri and Lev-On, 2007; Oprea et al., 2014), bargaining (e.g., Roth, 1995; Valley et al., 2002), and contract design (Brandts et al., 2016).[2] These studies typically focus on the effects of pre-play communication, i.e., how interacting parties change their actions when they are given the opportunity to send messages or talk to each other before making their choices. In contrast, we study how communication between humans and machines affects behavior while keeping the content of the communication constant across conditions. There are two notable exceptions that are more closely related to our study. Abeler et al. (2014) conducted an auxiliary experiment where subjects in the lab were asked to flip a coin four times and report their outcomes either via telephone or a computer interface. Conrads and Lotz (2015) additionally looked at the impact of face-to-face interaction. Consistent with our results, both studies find that the share of subjects who claim the maximum possible payoff decreases when reports are made by phone (or face-to-face) relative to a computer interface. However, reporting outcomes by phone or face-to-face differs from the computer interface not only by the presence of a human being but also in terms of how anonymous or identifiable subjects are in that situation.[3] Our study varies these two factors independently, effectively allowing us to isolate the role of social presence. For example, subjects experienced the same degree of anonymity in ROBOT and CALL as they reported their outcomes orally in both conditions. Understanding the independent role of human presence in honest behavior is important because when people interact with organizations, they rarely encounter the same degree of anonymity as they do in lab experiments. Although there might be some variation in identifiability across situations of interest, what mainly differs

[2] A few papers also explored the impact of non-binding promises on trust and trustworthiness (Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006; Vanberg, 2008; Corazzini et al., 2014).

[3] Several studies suggest that a lower degree of anonymity between parties (typically referred to as lower social distance) increases prosocial behavior (Hoffman et al., 1996; Bohnet and Frey, 1999; Charness et al., 2007; Charness and Gneezy, 2008).

is the presence of another person. Our study further departs from the existing literature by providing evidence that individuals anticipate the effect of social presence on their behavior and therefore avoid communication environments that require them to interact with another person.

Experiment 1 – Does digitized communication encourage dishonesty?

Design and procedures

We recruited subjects from the University of Zurich and the Swiss Federal Institute of Technology in Zurich (ETH) participant pool using the software h-root (Bock et al., 2014).[4] To recruit subjects, we first sent out an email eliciting their interest in participating in our study. Because the experiment was organized into individual sessions and required the software Skype, we asked potential subjects to indicate their availability and confirm that they had a Skype account. We also informed them that their data would be anonymized and treated confidentially and obtained their consent to participate in the study.[5] We then sent out a second email, asking subjects to select a time slot for their participation. We also reminded them that in order to participate, they would need a computer with stable Internet connection and Skype, and asked them to be in an undisturbed environment at the time of their participation.

At the beginning of a session, the experimenter contacted subjects on Skype to welcome them to the study. This stage was held constant across treatments to avoid differential selection based on initial contact. The experimenter first checked that the subjects were in a quiet place, and then told them that they need to get a piece of paper, a pen, and a coin. They then received a link to a short online survey that started with filler questions about life satisfaction and subjective well-being. Subsequently, subjects were instructed to flip a coin ten times and note the outcomes on paper. For each coin toss, they could earn 2 Swiss Francs (about US $2), depending on the outcome they reported at the end of the experiment. A payoff table indicated for each coin toss whether heads or tails would result in a monetary payoff (for more details on the procedures and instructions for the coin tossing task, see Online Appendix B). Subjects could increase their earnings by misreporting the outcomes of unsuccessful coin tosses. The stakes were significant, as subjects could earn up to 20 Swiss Francs within a relatively short amount of time (average completion time was about 14 minutes). Moreover, since subjects carried out the coin

[4]We excluded psychology students and subjects who had never participated in an economic lab experiment to ensure that they trusted our payment procedure. We also excluded individuals who had participated in previous experiments involving the coin tossing task.

[5]We obtained IRB approval from the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich.

tosses from a remote place, i.e., without being monitored, they could hide behind chance and nobody (including the experimenter) could determine with certainty whether a specific subject misreported their coin tosses. Subjects thus had a strong financial incentive to cheat without risk of getting caught.

However, reporting a high success rate might come across as suspicious and undermine a subject's appearance of being an honest person. Although it is impossible to identify cheating at the individual level, we are able to assess the extent of cheating in a group as the distribution from honest reporting is objectively defined. Specifically, assuming that none of the subjects cheated to their disadvantage (i.e., reporting that an outcome was not successful when it in fact it was so), we can estimate the cheating rates for the different conditions (see Houser et al. 2012): Let m be the percentage of misreported coin tosses. The percentage of outcomes reported as successful p is thus given by

$$p = m + (1 - m) \cdot 0.5 = 0.5 \cdot (1 + m). \quad (1)$$

A subject who cheats on a given coin toss reports a successful outcome with a probability of 1. In contrast, a subject who tells the truth reports a successful outcome only with a probability of 0.5. We can therefore characterize the percentage of misreported coin tosses as

$$m = 2 \cdot p - 1 \quad (2)$$

The coin tossing task (and variations of it) has been shown to reliably predict rule-violating behavior in natural settings, including violations of prison rules (Cohn et al., 2015), misbehavior in school (Cohn and Maréchal, 2016), absenteeism in the workplace (Hanna and Wang, 2014), free riding on public transport (Dai et al., 2016), and adulteration of milk (Kröll and Rustagi, 2016).

We implemented four treatments that varied, in a between-subjects design, how subjects reported the outcomes of their coin flips. In treatment CALL, subjects had to call the experimenter via Skype. They were instructed not to turn on the video feature. We did this to keep subjects' identifiability constant across conditions. In treatment FORM, subjects were provided with a link to a non-interactive online form where they could enter their outcomes. In treatment ROBOT, subjects used Skype to call an interactive voice response system with pre-recorded voice messages from the experimenters that asked them to report their outcomes. Thus, the only factor that changes compared to treatment CALL is whether or not subjects interacted with a real person. In particular, the degree of anonymity was held

Table 1. Overview of treatments in Experiment 1

	written communication	oral communication
machine interaction	FORM	ROBOT
human interaction	CHAT	CALL

constant because subjects reported orally in both treatments and therefore sent the same cues about their identity. Finally, subjects in treatment CHAT were asked to report their outcomes by chatting with the experimenter on Skype. We informed subjects about the communication channel before they tossed the coin. This is a complete 2x2 factorial design that varied (i) whether subjects interacted with another human being, and (ii) whether they reported their outcomes orally or in writing (see Table 1).

We designed our experiment so that the reporting of the coin flips was semantically identical across treatments. We used the same wording in each condition when we asked subjects to report their outcomes. They simply had to reply with "Heads" or "Tails" (either in writing or verbally) for each coin toss, permitting a ceteris paribus comparison of the treatments. In addition, we explicitly told all subjects that they would not be asked any questions in response to what they reported. Thus, they did not have to worry about justifying their reports in any of the treatments. Because some studies suggest that time pressure affects people's likelihood to cheat (e.g., Shalvi et al., 2012; Capraro, 2017; Lohse et al., 2018), we separated the actual tossing and reporting stages, and let subjects decide when they wanted to proceed to the reporting stage. This feature minimized differences in perceived time pressure between treatments.[6]

The experiment was conducted in two waves. The first took place in October and November 2013, the second one year later. The first wave included treatments CALL, FORM, and CHAT, while the second wave comprised treatments CALL, FORM, and ROBOT. A total of 486 subjects (n=257 in the first wave and n=211 in the second wave) participated in the first experiment. They were on average 24 years old and 49% were male. We employed five experimenters, two female and three male; two of them assisted us in each wave. Tables 1 and 2 in Online Appendix A present randomization checks and show that subjects' background characteristics are, with the exception of certain fields of study, well

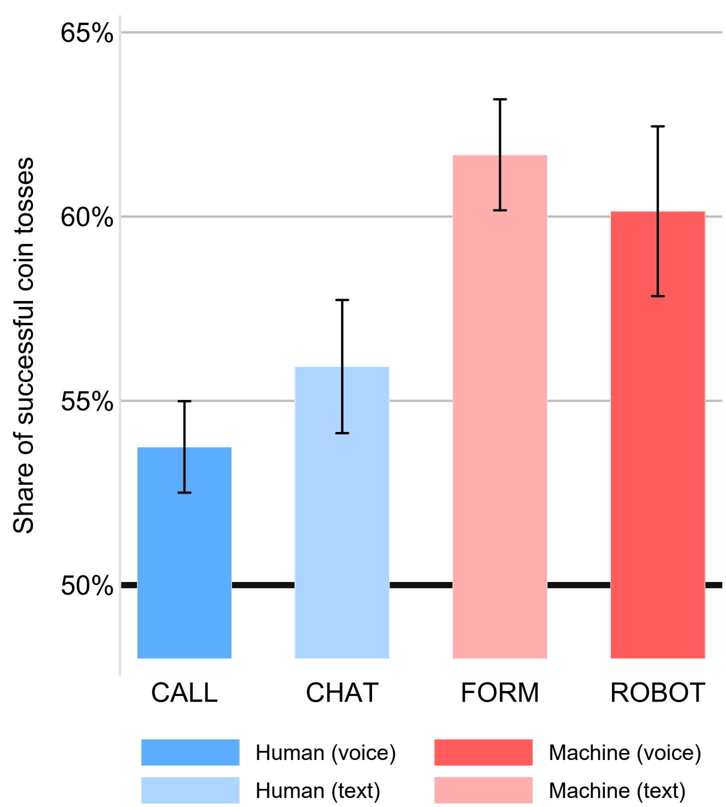
[6]We report supporting evidence in the last paragraph of the "Mechanism"-section. The time gap between the tossing and reporting stages also mimics many real-life situations, such as when a person involved in a car accident takes some time to process the event before calling the insurance company to report the damage.

balanced both across conditions and waves. The regression analyses control for subjects' field of study and other background characteristics.

Results

Figure 1 shows that subjects were relatively honest when calling the experimenter on Skype to report their coin flips. On average, they reported 53.8% successful coin flips in treatment CALL, which is only slightly though significantly higher than the honesty benchmark of 5.00% (95% confidence interval: 51.3–56.2%). This corresponds to a cheating rate of 8% (see equation 2). By contrast, cheating was more common when subjects used the online form. They reported 61.7% successful coin flips, on average, in treatment FORM (95% confidence interval: 58.7–64.7%), which is significantly higher than the average success rate reported in CALL ($p<0.001$, rank-sum test). The estimated cheating rate is 24% in FORM, and thus three times larger than in CALL. We replicate this effect if we split the data by waves. In the first wave, subjects reported 53.5% and 62.0% successful coin flips in CALL and FORM,

Figure 1. Information communication technology and cheating



Note: Percentage of successful coin tosses reported in each treatment condition. Error bars indicate standard error of the mean.

respectively (p=0.005, rank-sum test). The numbers are essentially the same in the second wave with success rates of 54.0% in CALL and 61.3% in FORM (p=0.035, rank-sum test).

Is the difference in cheating between CALL and FORM due to the richness of the communication medium (i.e., voice vs. text) or the presence of a real person? The results of treatment ROBOT suggest that real-time human interaction is essential for promoting honest behavior. Figure 1 shows that subjects cheated to a similar extent in ROBOT and FORM. They reported 60.1% successful coin tosses, on average, in ROBOT (95% confidence interval: 55.5–64.7%), which is not significantly different from FORM (p = 0.721, rank-sum test). Thus, subjects cheated to a similar degree when interacting with a machine, regardless of whether the machine was equipped with human cues. By contrast, subjects cheated significantly more in ROBOT than in CALL (p = 0.011, rank-sum test), despite both conditions being identical in terms of media richness and identifiability. Therefore, these results do not support social presence theory (Short et al., 1976), as the theory proposes that communication channels that transmit more human cues are more likely to activate peoples social image concerns. In other words, the vocal cues in treatment ROBOT seem unable to induce subjects to subconsciously believe that they were interacting with a human being.

We do not find evidence for social presence theory in the realm of human-human interaction either. Figure 1 shows that subjects reported 55.9% successful coin flips on average in treatment CHAT (95% confidence interval: 52.3-59.5%). Thus, while subjects reported more successful coin flips in CHAT than in CALL, the difference is small and not significant (p=0.363, rank-sum test). In contrast, the average success rate reported in CHAT is significantly lower than in FORM despite the fact that subjects used text to report their outcomes in both conditions (p=0.034, rank-sum test). Together, the results point to the importance of real-time human interaction rather than feature-richness of a communication medium for peoples tendency to cheat.[7]

We now turn to the regression analysis, which allows us to control for background characteristics. We estimate the following Probit model:

$$\Pr\left(y_{it} = 1 \mid \boldsymbol{T}_i, \boldsymbol{x}_i, \boldsymbol{z}_j\right) = \Phi\left(\alpha + \beta_1 \mathsf{FORM}_i + \beta_2 \mathsf{ROBOT}_i + \beta_3 \mathsf{CHAT}_i + \boldsymbol{\gamma}' \boldsymbol{x}_i + \boldsymbol{\delta}' \boldsymbol{z}_i\right) \tag{3}$$

[7] Media richness theory is another popular theory in communication research (Daft and Lengel, 1986). However, this theory makes the opposite prediction of social presence theory in regard to media richness and honesty. Specifically, the theory proposes that individuals will be more likely to cheat when using a communication medium that transmits more human cues, as it allows them to more effectively communicate complex and ambiguous information, such as a lie.

where $\Pr(\cdot)$ denotes the probability that subject i reported a successful outcome in trial t (i.e., $y_{it} = 1$), $\boldsymbol{T}_i$ represents a set of dummy of variables for treatments FORM, ROBOT, and CHAT (treatment CALL is therefore the reference category), $\boldsymbol{x}_i$ is a vector of individual background variables, including age, gender, Swiss nationality, and field of study (six categories), $\boldsymbol{z}_i$ are experimenter fixed effects, and Φ is the cumulative distribution function of the standard normal distribution. We report average marginal effects with standard errors clustered at the subject level to account for possible correlation of the residuals within individuals.

Table 2 presents the regression results, without (column 1) and with experimenter fixed effects (column 2). We find that subjects were 8 percentage points more likely to report a successful outcome in FORM than in CALL (p<0.001 in both columns, Wald tests; the base rate in CALL is 53.8%).

Table 2. Information communication technology and cheating

	(1)	(2)
Dependent variable	$y_{it} = 1$: coin toss reported as successful	
FORM	0.080***	0.080***
	(0.019)	(0.019)
ROBOT	0.063**	0.069***
	(0.025)	(0.025)
CHAT	0.020	0.017
	(0.021)	(0.022)
Age (years)	-0.001	-0.001
	(0.001)	(0.001)
Male subject	0.027*	0.027
	(0.017)	(0.017)
Swiss nationality	-0.015	-0.015
	(0.020)	(0.020)
Controls:		
Field of study	yes	yes
Experimenter FE	no	yes
Observations	4,680	4,680
Subjects	468	468

Note: Probit average marginal effect with standard errors, corrected for clustering at the individual level, displayed in parentheses. The dependent variable is always a dummy indicating whether a subject reported a coin toss as successful (10 observations per subject). The main independent variable are dummies which indicate whether a subject was in either treatments FORM, ROBOT, CHAT (CALL is omitted category). Additional independent variables include the subject's age in years, dummies for being male, Swiss citizenship, different fields of studies, and dummies to control for the experimenters' identities. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Likewise, subjects were about 7 percentage points more likely to report a successful outcome in ROBOT than in CALL (p=0.012 and 0.007, Wald tests), suggesting that the absence of a human counterpart encourages dishonest behavior. The difference between FORM and ROBOT is not significant (p = 0.540 and 0.683, Wald tests). Thus, making a machine more human by adding human voice has no impact on the interaction partner's honesty. In contrast, subjects in CHAT were about 6 percentage points less likely to report a successful outcome than those in FORM (p=0.010 and 0.007, Wald tests), but they were similarly likely to report a successful outcome relative to those in CALL (p=0.347 and 0.434, Wald tests). This suggests that media richness plays a minor role in honest behavior.

None of the experimenter fixed effects reach statistical significance at the conventional level, meaning that the experimenters did not affect the subjects differentially. Overall, the regression results confirm the preceding non-parametric analysis, suggesting that real-time human interaction promotes honest behavior.[8]

Mechanism

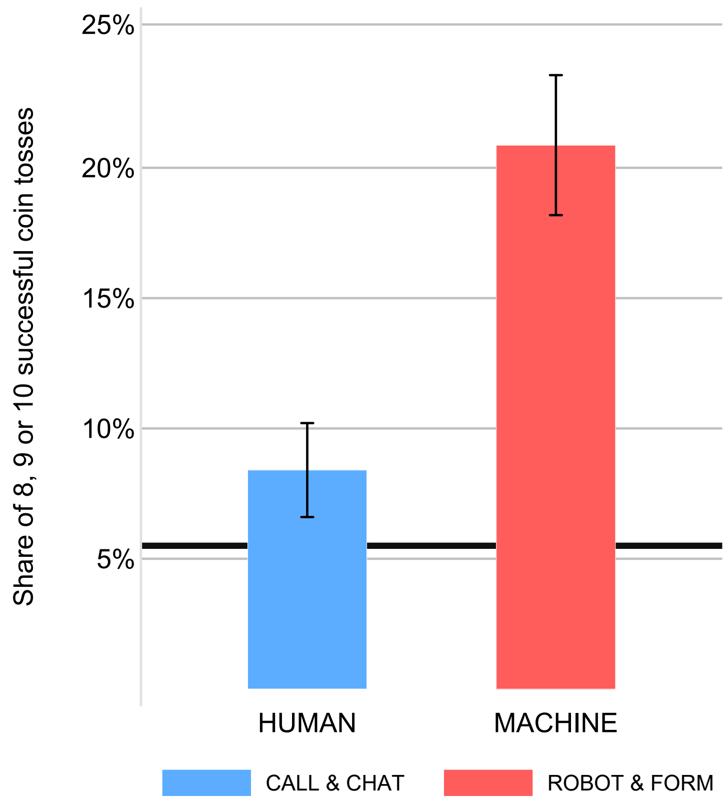
Why does human interaction encourage honest behavior? One possibility is that people care about their social image, i.e., they want to be perceived as honest (Abeler et al., 2016; Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2016; Khalmetski and Sliwka, 2017; see also Bursztyn and Jensen, 2017 for a recent review of the social image literature). However, individuals may pay less attention to their image when interacting with a machine relative to a human being. If true, we should observe that subjects are less likely to report high, and therefore potentially suspicious, success rates such as 8, 9, or 10 successful coin flips less often to a human (i.e., treatments CALL or CHAT) than to a machine (i.e., treatments FORM or ROBOT).[9]

Figure 2 is very much in line with this prediction. We find that 8.4% (95% confidence interval: 4.9–12.0%) of the subjects reported 8 or more successful coin flips when they report to a human being. Given that we should expect 5.5% of the subjects to do so if everyone reports truthfully, this suggests that only a few more individuals than predicted by chance claimed a high success rate when reporting to the experimenter. In sharp contrast, the share of subjects reporting 8, 9 and 10 winning coin flips is 20.9% (95% confidence interval: 15.6–26.2%) when they reported to a machine. Thus, the excess

[8]We further find that students in business and economics are more likely to cheat (p=0.018 and 0.017, Wald tests). This is line with previous research that documents a positive correlation between studying economics and self-interested behavior (see Frank et al., 1993; Frey and Meier, 2003; López-Pérez and Spiegelman, 2012; Cappelen et al., 2015)

[9]The probability of 8, 9, and 10 successful coin flips is 4.4%, 1.0%, and 0.1%, respectively. For comparison, the probability of 6 and 7 successful coin flips is 20.5% and 11.7%, respectively.

Figure 2. Extreme reports in human versus machine interaction



Note: Percentage of subjects who reported 8 or more successful coin tosses for interactions with humans (i.e., treatments CALL and CHAT) and machines (i.e., treatments FORM and ROBOT). Error bars indicate standard error of the mean.

incidence of unlikely high success rates is more than five times larger in the machine relative to the human conditions, suggesting that subjects felt more at ease to cheat outright when they reported to a machine.

We also estimate the following Probit model to underpin these results statistically:

$$\Pr\left(y_i \in \{8, 9, 10\} \mid \mathsf{MACHINE}_i, \boldsymbol{x}_i, \boldsymbol{z}_j\right) = \Phi\left(\alpha + \beta_1 \mathsf{MACHINE}_i + \boldsymbol{\gamma}'\boldsymbol{x}_i + \boldsymbol{\delta}'\boldsymbol{z}_i\right) \tag{4}$$

where $\Pr(\cdot)$ denotes the probability that subject i reported 8, 9, or 10 successful coin flips. $\mathsf{MACHINE}_i$ is an indicator which takes a value of one if the subject reported to a machine (i.e., treatment FORM or ROBOT). As before, $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ capture control variables and experimenter fixed effects. We report average marginal effects with robust standard errors. As a robustness check, we also estimate the same model using a dummy for more plausible success rates as dependent variable (i.e., 6 or 7 successful coin flips).

Column (1) in Table 3 reports the results for high and therefore unlikely success rates (i.e., 8 or more successful coin flips). We find that the share of subjects reporting 8, 9, or 10 successful coin flips is 15.4 percentage points higher in the machine conditions, which is statistically significant ($p<0.001$, Wald test).[10] We do not observe such a pattern for more plausible success rates (i.e., 6 or 7 successful coin flips), as shown in column (2). Although subjects reported 6 or 7 successful coin flips more often than predicted by chance (40.7% instead of 32.2%; 95% confidence interval: 34.4–47.0%), they did so to a similar extent regardless of whether they reported to a person or a machine ($p=0.862$, Wald test). Together, these results suggest that subjects were reluctant to report success rates that are unlikely and

Table 3. Human interaction and unlikely outcomes

	(1)	(2)
Dependent variable:	$y_i \in \{8, 9, 10\}$	$y_i \in \{6, 7\}$
MACHINE	0.154***	-0.008
	(0.032)	(0.046)
Base rate	0.084***	0.407***
	(0.018)	(0.031)
Expected rate	0.055	0.322
Controls:		
Age	yes	yes
Gender	yes	yes
Nationality	yes	yes
Field of study	yes	yes
Experimenter FE	yes	yes
Observations	468	468

Note: Probit average marginal with robust standard errors displayed in parentheses. The dependent variable in column (1) is a dummy indicating whether a subject reported 8, 9, or 10 successful coin tosses. In column (2), the dependent variable indicates whether a subject reported 6 or 7 successful coin tosses. The main independent variable MACHINE is a dummy which indicates whether a subject reported to a machine (i.e., treatments FORM or ROBOT). The two treatments with human interaction CALL and CHAT are omitted and serve as the reference category. Base rate refers to the average of the respective dependent variables across these treatments (i.e., CALL and FORM). Expected rate is the respective base rate one would expect if everyone reported truthfully. Control variables include the subject's age in years and dummies for being male, Swiss citizenship, different fields of studies, and the experimenters' identities. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

[10] We obtain qualitatively similar results if we take a cutoff at 9 instead of 8 successful coin flips or if we do not use the pooled MACHINE but the individual treatment dummies.

may therefore raise suspicion when interacting with a person. This is consistent with the notion that people have a desire to maintain a positive social image.

We also explore alternative explanations for why subjects cheated less when interacting with a human rather than a machine. For example, while the detection probability was effectively zero in all conditions, it is nevertheless conceivable that some subjects erroneously thought that they could get caught cheating and that they would not get paid. If true, we should observe that our results differ depending on subjects' risk attitudes. Specifically, we should see that (i) more risk-averse subjects were generally less likely to cheat, and (ii) that the difference between the machine and human conditions is larger for risk-averse subjects, as they might have worried more about getting caught when they interacted with the experimenter.

To address this, we elicited subjects' risk attitudes using an experimentally validated survey question developed by Dohmen et al. (2011). Specifically, we asked subjects "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks" using an 11-point Likert scale ranging from "not at all willing to take risk" to "very willing to take risks." For ease of interpretation, we reversed coded the answers and normalized the values to a mean of zero and a standard deviation of one. The resulting variable can thus be interpreted as a proxy for risk aversion in standard deviation units. We then estimate a Probit-model very similar to (4) which differs in two aspects: First, we use whether coin toss t by subject i is reported as successful ($y_{it} = 1$) as the dependent variable. Second, we add our measure for risk aversion and its interaction with the MACHINE-dummy as independent variables. This interaction term allows testing whether more risk-averse subjects are more sensitive to the presence or absence of another person to which they report.

Table 4 presents the results in three steps. Column (1), which presents the results without controlling for our measure of risk aversion, shows that subjects were about 7.2 percentage points more likely to report a successful outcome when reporting to a machine rather than a person ($p<0.001$, Wald test). Column (2) indicates that, across conditions, subjects' risk aversion does not significantly predict their likelihood of reporting successful outcomes ($p=0.756$, Wald test). This suggests that our experimental design successfully eliminated perceived threats of punishment. Moreover, controlling for risk aversion does not change the coefficient of the MACHINE-dummy. Column (3) reports the results when we additionally include the interaction between risk aversion and the MACHINE-dummy. This allows us to estimate the slope of the risk aversion-coefficient separately for the machine and human conditions. If subjects thought that the chance of getting caught and punished was smaller when reporting to a machine, we should see that more risk-averse individuals felt significantly more comfortable cheating in the machine conditions. However, we find that the coefficient of risk aversion remains small and is not significantly different from zero in either condition (p=0.192 and p=0.428, Wald test). Also, the difference between the two coefficients is not statistically significant (p=0.143, Wald-Test). Overall, the results do not support the conjecture that punishment concerns drive our main result.

Could our findings be explained by differences in time pressure across conditions? Perhaps subjects felt more pressure when reporting to the experimenter, and consequently cheated less (see Capraro, 2017; Lohse et al., 2018). We minimized this possibility in two ways. First, we informed subjects

Table 4. Risk aversion and cheating

	(1)	(2)	(3)
Dependent variable	$y_{it} = 1$: coin toss reported as successful		
MACHINE	0.072***	0.072***	0.073***
	(0.016)	(0.016)	(0.016)
Risk aversion		-0.003	
		(0.009)	
Risk aversion (at MACHINE=0)			-0.015
			(0.012)
Risk aversion (at MACHINE=1)			0.010
			(0.013)
Controls			
Age	yes	yes	yes
Gender	yes	yes	yes
Nationality	yes	yes	yes
Field of Study	yes	yes	yes
Experimenter FE	yes	yes	yes
Observations	4,680	4,680	4,680
Subjects	468	468	468

Note: Probit average marginal effect with standard errors, corrected for clustering at the individual level, displayed in parentheses. The dependent variable is a dummy indicating whether subjects reported a coin toss as successful (10 observations per subject). The main independent variable MACHINE is a dummy which indicates whether a subject reported to a machine (i.e., treatments FORM or ROBOT). The two treatments with human interaction CALL and CHAT are omitted and serve as the reference category. Column (2) includes a proxy for subjects' risk aversion based on their response to the question "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks" using an 11-point Likert scale ranging from "not at all willing to take risk" to "very willing to take risks." We normalized the risk aversion proxy to a mean of zero and standard deviation of one. Column (3) displays the relationship between risk aversion for the two values of the MACHINE-dummy. Control variables include the subject's age and dummies for being male, Swiss citizenship, different fields of studies, and the experimenters' identities. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

that they would not be asked any questions about what they report. Second, we separated the coin tossing and reporting stage to give subjects enough time to think about what they want to report. To further investigate the plausibility of this explanation, we also elicited subjects' perceived time pressure by asking them "To what extent did you feel under time pressure when reporting the outcomes of your coin tosses?" using a 7-point Likert scale with possible answers ranging from "not at all" (=0) to "very much" (=6). Yet, 78% of the subjects reported a zero or one on this scale, indicating that a majority of the subjects did not feel any pressure when they reported their outcomes. Perceived time pressure is not only low across all conditions, but also almost identical between the human and machine conditions (0.90 vs. 0.91, p=0.625, rank-sum test). Thus, it is unlikely that subjects were cheating less in the human conditions because of time pressure.

Experiment 2 – Do dishonest people prefer interacting with a machine?

The results of the first experiment suggest that individuals behave more dishonestly when they interact with a machine compared to a human being. Could self-selection into communication channels be used as a device to screen for dishonest people? To find out, we conducted a second experiment in which subjects could choose between reporting their coin flips to the experimenter or to a machine. If dishonest individuals anticipate that they will feel less comfortable misreporting unsuccessful coin flips to a person, they should prefer to report their outcomes to the machine.

Design and procedures

We recruited a new sample of subjects for the second experiment using a similar procedure as for the first experiment. In the invitation email, we additionally explained that the study consists of two parts, taking place roughly one week apart. While subjects completed the first part (Part A) independently, they had to indicate their availability for the second part (Part B) so that we could schedule individual sessions with an experimenter. We further told them that, although they had to sign up and commit to participate in both parts, only every fourth subject, selected at random, would eventually participate in Part B.[11] Those selected to participate in both parts were paid according to their responses in one of the two parts, which was randomly determined at the end of the study. We chose this procedure to prevent carry-over effects between the two parts. Subjects selected to participate only in Part A were

[11]We limited the number of participants for Part B because our main focus is subjects' choices in Part A.

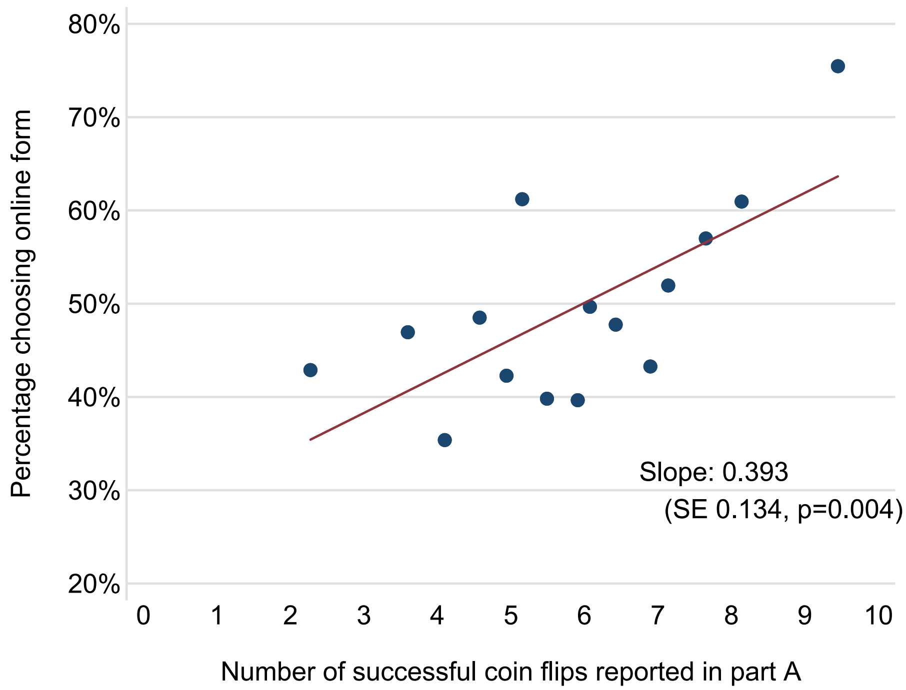
paid based on their responses in that part. Finally, we assured participants that their data would be anonymized and treated confidentially, and obtained their informed consent.

Subjects then received an email on a pre-announced date which asked them to complete a short online survey by the end of the day. The survey began with the same filler questions about life satisfaction and subjective well-being as in the first experiment. Just like in treatment FORM, subjects were subsequently instructed to perform ten coin tosses and to report the outcomes online using a non-interactive form. Each coin toss could yield a payoff of 2 Swiss Francs. Because higher earnings are less likely to be the result of chance, we can use the earnings from this coin tossing task as a proxy for individuals' tendency to cheat. At the end of the survey, subjects were instructed to toss the coin another ten times and note the outcomes on paper. Then, they were asked to choose how to report the outcomes of the second coin tossing task in Part B of the study. They could choose between reporting their results to the experimenter via Skype call (without video), or they could use the online form. The two options were presented in randomized order. Those subjects who were selected for participation in Part B received a Skype call from the experimenter a few days later at the agreed date and time. Thus, Part B always started with a quick Skype call, regardless of whether subjects chose to report their coin flips using the online form (and subjects knew this at the time of their choice). They were then either sent a link to the online form or reported their outcomes to the experimenter, depending on the choice they made in Part A (see instructions in Online Appendix C). A total of 380 subjects participated in this experiment. They were 23 years old, on average, and 47.4% were male (see Table 3 in Online Appendix A). We employed one experimenter for Part B.

Results

Overall, subjects were equally likely to choose either of the two communication channels. 50.5% of the subjects chose to call the experimenter on Skype, and 49.5% of them chose the online form to report their outcomes of the second coin tossing task ($p=0.877$, two-sided binomial test with 50% as null hypothesis). This suggests that both ways of reporting were perceived as similarly convenient. However, the binned scatter plot (following the procedure of Chetty et al. 2014) shows that subjects who report a higher number of successful coin flips in part A of the experiment (i.e., those who have presumably cheated) were also more likely to choose the online form to report their outcomes of the second coin tossing task in part B.

Figure 3. Screening for presumably dishonest people



Note: Binned scatter plot (following the procedure of Chetty et al. 2014) illustrating the relationship between the number of successful coin tosses in part A and the likelihood of choosing to report the outcomes of the subsequent coin tossing task through the online form in part B. To construct this plot, we regress both the choice to use the online form (y-axis variable) and the number of successful coin flips (x-axis variable) on our standard set of controls using OLS and calculate the residuals. We then group the residuals of the x-axis variable into fifteen equally-sized bins. Within each bin, we compute the mean of the x- and y-axis' residuals and add the respective variable's unconditional sample mean to create a scatterplot of these data points. The solid line represents the OLS regression line based on the underlying individual data.

We corroborate these results by estimating a Probit model of the following form:

$$\Pr\left(c_i = \text{FORM} \mid y_i^A, \boldsymbol{x}_i\right) = \Phi\left(\alpha + \beta_1 y_i^A + \boldsymbol{\gamma}'\boldsymbol{x}_i\right) \tag{5}$$

where $\Pr(\cdot)$ denotes the probability that subject i selected the online form for reporting the second set of coin tosses, y_i^A is the number of successful coin flips from the first coin tossing task, and $\boldsymbol{x}_i$ is our standard set of control variables for subjects' background characteristics. We report average marginal effects with robust standard errors.

Table 5 presents the estimation results. For every successful coin flip in Part A, subjects were about 3.9 percentage points more likely to select the online form for Part B (p=0.005 and p=0.003, Wald

Table 5. Selection into machine reporting

	(1)	(2)
Dependent variable	c_i = FORM: choice to report to a machine	
Successful coin tosses in part A	0.038***	0.039***
	(0.013)	(0.013)
Controls:		
Age	no	yes
Gender	no	yes
Nationality	no	yes
Field of study	no	yes
Observations	380	380

Note: Probit average marginal effect with robust standard errors displayed in parentheses. The dependent variable is a dummy variable indicating whether a subject chose to report to a machine in part B of experiment 2. The main independent variable is the number of successful coin tosses stated in part A. Additional independent variables in column (2) are the subject's age in years and dummies for being male, Swiss citizenship, and different fields of studies. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

tests). In other words, a subject who claimed to have been successful on every possible instance is 19 percentage points more likely to choose the online form relative to someone who reported 50% successful coin flips. This results remains the same when we control for subjects' background characteristics. By contrast, none of the subjects' background characteristics (i.e., age, gender, nationality, and field of study) predicts their choices of communication channel significantly. In sum, when given the choice, alleged cheaters avoid human interaction to report private information that they can manipulate to their own material benefit.[12]

Conclusion

The digital age has radically changed the way we communicate and interact with each other. For example, we walked over to the local branch of the insurance company to report a stolen bicycle 50 years ago, we called a representative of the insurance company 20 years ago, and we can just fill out an online form or interact with a chatbot insurance agent today. Are we more likely to misrepresent information when submitting an insurance claim online rather than in person or over the phone? In this paper, we examine the importance of human interaction in digital communication when individuals have an incentive to exploit informational asymmetries to their own advantage. Our experimental paradigm

[12] In the second part of the study (with a reduced sample), we find that subjects were 8.5 percentage points more likely to report a successful outcome when choosing the online form relative to calling the experimenter on Skype (p=0.009, Wald test).

for measuring dishonest behavior is a coin tossing task in which subjects are asked to privately flip a coin multiple times, report the outcomes from those coin flips, and then receive a payment depending on the outcomes they report.

In the first experiment, we varied the communication channel through which subjects had to report their coin flips and found that they cheat substantially less when they interact with a person rather than a machine. Human interaction appears to be essential for honest reporting because equipping a machine with human features (i.e., a human voice) does not encourage more honest behavior. Further analysis proposes a mechanism based on social image concerns, i.e., individuals want to maintain an honest appearance towards others, even if they will never meet them again. This underscores the importance of real-time human interaction in reducing fraudulent behavior, a finding that has implications for organizations which rely on customers' or employees' willingness to behave honestly, such as banks, insurance companies, or tax collection authorities. But, of course, employing people is costly and may not necessarily offset the benefits of reduced fraud. Nonetheless, our study ascribes a powerful role to human interaction in mitigating dishonest behavior and therefore speaks to growing concerns that robots and other computer-assisted technologies might render many of today's workers obsolete (e.g., Autor, 2015; Acemoglu and Restrepo, 2017).

We conducted a second experiment to examine whether individuals with a greater tendency to cheat avoid communication channels that require interacting with human beings. We indeed find that subjects who are more likely to cheat prefer reporting their coin flips to a machine. This raises the possibility for companies to screen for customers with an increased propensity to engage in fraudulent activities. For example, firms could offer customer service through multiple communication channels that differ by whether customers interact with a real agent. The self-selection of customers into different communication channels is a promising tool for targeting monitoring efforts toward cases with high risk of fraud and thereby reducing the cost of fraud detection.

References

Abeler, J., A. Becker, and A. Falk (2014). Representative Evidence on Lying Costs. *Journal of Public Economics 113*, 96–104.

Abeler, J., C. Raymond, and D. Nosenzo (2016). Preferences for Truth-Telling. *IZA Discussion Paper 10188*.

Acemoglu, D. and P. Restrepo (2017). Robots and Jobs: Evidence from US Labor Markets. *mimeo*.

Andreoni, J. and D. Bernheim (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica 77*(5), 1607–1636.

Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*(1), 544–555.

Autor, D. H. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives 29*(3), 3–30.

Bateson, M., D. Nettle, and G. Roberts (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters 2*(3), 412–414.

Bicchieri, C. and A. Lev-On (2007). Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. *Politics, Philosophy & Economics 6*(2), 139–168.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot - Hamburg registration and organization online tool. *European Economic Review 71*(117-120).

Bohnet, I. and B. S. Frey (1999). Social Distance and other-regarding Behavior in Dictator Games: Comment. *American Economic Review 89*(1), 335–339.

Brandts, J., M. Ellman, and G. Charness (2016). Let's Talk: How Communication Affects Contract Design. *Journal of the European Economic Association 14*(4), 943–974.

Brewster, S. (2016). Do Your Banking with a Chatbot. *MIT Technology Review* (May 17, 2016).

Bureau of Labor Statistics (2016). 24 percent of employed people did some or all of their work at home in 2015. *U.S. Department of Labor: The Economics Daily* (July 8, 2016).

Bursztyn, L., B. Ferman, S. Fiorin, M. Kanz, and G. Rao (2017). Status Goods: Experimental Evidence from Platinum Credit Cards. *NBER Working Paper Series*, 64.

Bursztyn, L. and R. Jensen (2017). Social Image and Economic Behavior in the Field: Identifying, Understanding and Shaping Social Pressure. *Annual Review of Economics 9*, 131–153.

Cai, W., X. Huang, S. Wu, and Y. Kou (2015). Dishonest behavior is not affected by an image of watching eyes. *Evolution and Human Behavior 36*(2), 110–116.

Cappelen, A. W., K. Nygaard, E. Sørensen, and B. Tungodden (2015). Social Preferences in the Lab: A Comparison of Students and a Representative Population. *Scandinavian Journal of Economics 117*(4), 1306–1326.

Capraro, V. (2017). Does the truth come naturally? Time pressure increases honesty in one-shot deception games. *Economics Letters 158*, 54–57.

Charness, G. and M. Dufwenberg (2006). Promises and Partnership. *Econometrica 74*(6), 1579–1601.

Charness, G. and U. Gneezy (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization 68*, 29–35.

Charness, G., E. Haruvy, and D. Sonsino (2007). Social distance and reciprocity: An Internet experiment. *Journal of Economic Behavior and Organization 63*(1), 88–103.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Cohn, A. and M. A. Maréchal (2016). Laboratory Measure of Cheating Predicts Misbehavior at School. *Economic Journal*, forthcoming.

Cohn, A., M. A. Maréchal, and T. Noll (2015). Bad Boys: How Criminal Identity Salience Affects Rule Violation. *Review of Economic Studies 82*(4), 1289–1308.

Conrads, J. and S. Lotz (2015). The effect of communication channels on dishonest behavior. *Journal of Behavioral and Experimental Economics 58*, 88–93.

Cooper, R., D. V. Dejong, R. Forsythe, and T. W. Ross (1992). Communication in Coordination Games. *Quarterly Journal of Economics 107*(2), 739–771.

Corazzini, L., S. Kube, M. A. Maréchal, and A. Nicolò (2014). Elections and Deceptions: An Experimental Study on the Behavioral Effects of Democracy. *American Journal of Political Science 58*(421), 579–592.

Crawford, V. (1998). A Survey of Experiments on Communication via Cheap Talk. *Journal of Economic Theory 78*(2), 286–298.

Daft, R. L. and R. H. Lengel (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science 32*(5), 554 – 571.

Dai, Z., F. Galeotti, and M. C. Villeval (2016). Cheating in the Lab Predicts Fraud in the Field. An Experiment in Public Transportations. *Management Science*, forthcoming.

Dawkins, R. (2006). *The God Delusion*. New York: Houghton Mifflin Harcourt.

DellaVigna, S., J. A. List, and U. Malmendier (2012). Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics 127*(1), 1–56.

DellaVigna, S., J. A. List, U. Malmendier, and G. Rao (2017). Voting to Tell Others. *Review of Economic Studies 84*(1), 143–181.

Dufwenberg, M. and M. A. Dufwenberg (2018). Lies in Disguise - A Theoretical Analysis of Cheating. *Journal of Economic Theory 175*, 248–264.

Ellingsen, T. and M. Johannesson (2004). Promises , Threats and Fairness. *Economic Journal 114*(495), 397–420.

Ernest-Jones, M., D. Nettle, and M. Bateson (2011). Effects of eye images on everyday cooperative behavior: A field experiment. *Evolution and Human Behavior 32*(3), 172–178.

Fehr, E. and F. Schneider (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences 277*(1686), 1315–1323.

Frank, R. H., T. Gilovich, and D. T. Regan (1993). Does Studying Economics Inhibit Cooperation? *Journal of Economic Perspectives 7*(2), 159–171.

Frey, B. S. and S. Meier (2003). Are political economists selfish and indoctrinated? Evidence from a natural experiment. *Economic Inquiry 41*(3), 448–462.

Gächter, S. and J. F. Schulz (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature 531*(7595), 496–499.

Gneezy, U. (2005). Deception: The Role of Consequences. *American Economic Review 95*(1), 384–394.

Gneezy, U., A. Kajackaite, and J. Sobel (2016). Lying Aversion and the Size of the Lie. *American Economic Review*, forthcoming.

Haley, K. J. and D. M. T. Fessler (2005). Nobody's watching? Subtle cues affect generosity an anonymous economic game. *Evolution and Human Behavior 26*, 245–256.

Hall, S. (2017). How Artificial Intelligence Is Changing The Insurance Industry. *National Association for Insurance Policy and Research - CIPR Newsletter 22*.

Hanna, R. and S.-Y. Wang (2014). Dishonesty and selection into public service. *American Economic Journal: Applied Economics*, forthcoming.

Hoffman, B. E., K. Mccabe, and V. L. Smith (1996). Social Distance and Other-Regarding Behavior in Dictator Games. *American Economic Review 86*(3), 653–660.

Hortaçsu, A. and C. Syverson (2015). The Ongoing Evolution of US Retail: A Format Tug-of-War. *Journal of Economic Perspectives 29*(4), 89–112.

Irlenbusch, B. and M. C. Villeval (2015). Behavioral ethics: How psychology influenced economics and how economics might inform psychology? *Current Opinion in Psychology 6*, 87–92.

Isaac, R. M. and J. M. Walker (1988). Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism. *Economic Inquiry 26*(4), 585–608.

Khalmetski, K. and D. Sliwka (2017). Disguising Lies - Image Concerns and Partial Lying in Cheating Games. *CESIFO Working Paper 6347*.

Kosfeld, M. and S. Neckermann (2011). More Work for Nothing? Getting Symbolic Awards and Worker. *American Economic Journal: Microeconomics 3*(3), 86–99.

Kröll, M. and D. Rustagi (2016). Shades of dishonesty and cheating in informal milk markets in India. *SAFE Working Paper 134*.

Kurzban, R., P. DeScioli, and E. O'Brien (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior 28*(2), 75–84.

Lohse, T., S. A. Simon, and K. A. Konrad (2018). Deception under time pressure: Conscious decision or a problem of awareness? *Journal of Economic Behavior & Organization 146*, 31–42.

López-Pérez, R. and E. Spiegelman (2012). Do Economists Lie More? *mimeo*.

Mateyka, B. P. J., M. A. Rapino, and L. C. Landivar (2012). Home-Based Workers in the United States: 2010. *U.S. Department of Commerce - Current Population Reports*.

Mazar, N., O. Amir, and D. Ariely (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research 45*(6), 633–644.

Northover, S. B., W. C. Pedersen, A. B. Cohen, and P. W. Andrews (2017). Artificial surveillance cues do not increase generosity: two meta-analyses. *Evolution and Human Behavior 38*(1), 144–153.

Oprea, R., G. Charness, and D. Friedman (2014). Continuous time and communication in a public-goods experiment. *Journal of Economic Behavior and Organization 108*, 212–223.

Roth, A. E. (1995). Bargaining experiments. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, pp. 253–348. Princeton: Princeton University Press.

Shalvi, S., O. Eldar, and Y. Bereby-Meyer (2012). Honesty Requires Time (and Lack of Justifications). *Psychological Science 23*, 1264–1270.

Shalvi, S., F. Gino, R. Barkan, and S. Ayal (2015). Self-Serving Justifications. *Current Directions in Psychological Science 24*(2), 125–130.

Short, J., E. Williams, and B. Christie (1976). *The social psychology of telecommunications*. London: Wiley.

Trivers, R. (2006). Reciprocal altruism: 30 years later. In P. M. Kappeler and C. P. van Schaik (Eds.), *Cooperation in primates and humans: Mechanisms and evolution*. New York.

Turkle, S. (2012). The flight from conversation. *The New York Times* (April 21, 2012).

U.S. Department of Commerce (2017). Quarterly Retail E-Commerce Sales - 2nd Quarter 2017. *U.S. Census Bureau News* (August 17, 2017).

Valley, K., L. Thompson, R. Gibbons, and M. H. Bazerman (2002). How communication improves efficiency in bargaining games. *Games and Economic Behavior 38*, 127–155.

Vanberg, C. (2008). Why Do People Keep Their Promises? An Experimental Test of Two Explanations. *Econometrica 76*(6), 1467–1480.