Honesty in the Digital Age

October 2020

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 a total of 848 subjects to examine how human versus machine interaction influences cheating for financial gain. We find that individuals cheat about three times 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 and, therefore, suspicious 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 presence is key to mitigating dishonest behavior and that self-selection into communication channels can be used to screen for dishonest people.

Keywords: Honesty, cheating, human interaction, digitization, social image, self-selection
JEL Classification: C99, D82, D83

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@ethz.ch, Center for Law & Economics, ETH Zurich
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 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 concerns, i.e., they care about what others think of them (see Bursztyn and Jensen, 2017, for a recent overview). In the absence of a human interaction partner, individuals may care less about leaving a good impression on others. As a result, they might feel more comfortable lying to a machine than a person.[1] What if the machine is made more human, i.e., it is capable of mimicking (at least to some degree) a real person? For example, it is possible that the interaction with a machine that has human features might make people feel like they are interacting with a real person, which in turn may trigger similar image concerns.

We examine these questions 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 then paid them according

[1] Media richness theory (Daft and Lengel, 1986), a popular theory in communication research, makes the opposite prediction. This 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 communicate complex and ambiguous information, such as a lie, more persuasively. For example, in a diary-based study, Hancock et al. (2004) asked people how often they lie across different communication channels. They found that people reported being twice as likely to lie in face-to-face conversations as when communicating by email.

to their alleged success rates. Subjects performed the coin tosses from a remote place (typically from home) and reported their outcomes via the communication software Skype. We varied, across two waves of data collection, (i) whether subjects reported their outcomes to a person or a machine, and (ii) whether the interaction involved oral or written communication. The first wave comprised three conditions. 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 keep the degree of anonymity constant across conditions. 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 present in the reporting stage in treatment FORM. However, because the two conditions also varied the communication mode (oral vs. written), we implemented a third condition, treatment CHAT. In this condition, subjects had to report their outcomes to the experimenter in writing (using Skype instant messaging). The comparison of treatments CALL and CHAT allows us to test whether the transmission of additional human cues (i.e., voice instead of just text) affects people's tendency to cheat in human interactions. The extent of human cues could be relevant in light of the literature on computer-mediated communication studying the role of verbal and nonverbal cues in deception detection across online and offline contexts (e.g., for reviews see Hancock and Guillory, 2015; Toma et al., 2019). For the second wave, we replicated our main treatments CALL and FORM, and also introduced a new treatment called ROBOT. Treatment ROBOT was identical to treatment CALL, except that subjects communicated with an automated voice response system that prompted them to report their outcomes using pre-recorded voice messages of the experimenters. Thus, relative to treatment CALL, we only varied the presence of a real person at the time of the reporting while keeping the communication mode constant across conditions.

Several design features are noteworthy. First, we designed the experiment so that in all conditions subjects used exactly the same wording (either orally or in writing) to report their outcomes. This allows for a *ceteris paribus* comparison of the treatments. We further informed subjects that they will not be asked any follow-up questions about what they report. Thus, subjects did not have to worry about justifying a high number of successful coin flips – even when they interacted with a person. Second, because subjects performed the coin-tossing task in private, there was no way the experimenters (nor anyone else) could unambiguously identify whether a specific subject cheated. Thus, the risk of getting punished for cheating was zero in all treatments. Finally, we gave subjects enough time in all conditions to perform the coin flips and think about how many coin flips they wanted to report as successful before

they moved on to the reporting stage. Thus, any differences between conditions cannot be attributed to varying time pressure, a factor that has been previously shown to affect cheating (e.g., Shalvi et al., 2012; Capraro, 2017; Lohse et al., 2018).

Across both waves, we find that subjects reported a higher success rate in treatment CALL compared to FORM. They reported 53.5% (54.0%) successful coin flips, on average, in CALL of Wave 1 (Wave 2). This corresponds to a cheating rate of 7.6% if we pool both waves and assume that none of the subjects cheated to their disadvantage. By contrast, subjects reported 62.0% (61.3%) successful coin flips, on average, in FORM of Wave 1 (Wave 2). This corresponds to a cheating rate of 23.4% in FORM across both waves, which is about three times as large as in CALL. Changing the communication mode (i.e., voice vs. text) did not affect the level of cheating by much. In treatment CHAT, subjects reported 55.9% successful coin flips, on average, and were therefore similarly honest as those in CALL. By contrast, replacing human with machine interaction had a considerable impact on cheating. In treatment ROBOT, subjects reported 60.1% successful coin flips, on average. Thus, subjects cheated to a similar degree as in FORM, despite the fact that they could hear the experimenter's voice, which presumably made the machine interaction more human.

A potential mechanism for these results is that people care less about their social image (i.e., being judged as honest) when they interact with a machine rather than a human. Although in our experiment there was always a human experimenter involved at some point (e.g., when welcoming participants or for making the payments), the key difference across human and machine conditions is the variation in the presence of a human being during the reporting stage. Hence, our experiment manipulated subjects' sense of human presence—i.e., the feeling of closeness in terms of socially interacting with another person.[2] A stronger sense of human presence may enhance subjects' concerns about being judged and what the experimenter thinks of them. Indeed, an additional survey experiment indicates that, relative to human interaction, reporting to a machine reduces people's feeling of another person being present and it also diminishes their social image concerns. Moreover, in our main experiment 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 compared to a person. By contrast, more credible success rates (i.e., 6 or 7 successful coin tosses) were reported with similar frequency across

[2]We use the term "human presence" rather than "social distance" to avoid confusion about terminology. In prior research, the notion of social distance has been conceptualized primarily in two ways: (i) the degree of subject anonymity (e.g., Hoffman et al., 1996) and (ii) the extent to which people are similar and share the same social identity (e.g., Akerlof, 1997). In the section "Alternative explanations," we discuss why our results cannot be explained by anonymity or similarity considerations.

human and machine conditions. This suggests that the treatment differences in honest behavior can be explained by different levels of social image concerns.

The results of this experiment raise the question of whether dishonest people anticipate that they would feel more uncomfortable misrepresenting information when they interact with a person compared to 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. In this first coin tossing task all subjects reported their outcomes under identical conditions.

When asked to choose between communication channels, subjects were about equally likely to select the call and the online form (50.5% vs. 49.5%). 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 for each additional successful coin flip in the first part, subjects were about 4 percentage points more likely to choose the online form for the second part. Thus, more dishonest individuals avoid human interaction when they have an opportunity to cheat for personal financial gain. This "selection on moral hazard" raises the possibility for organizations to screen for dishonest customers. For example, firms could increase the effectiveness of their auditing activities by offering clients the choice between different communication channels and targeting those who choose not to interact with a live customer representative.

Our paper relates to several strands of the literature. First, our findings contribute to a growing literature 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 honest behavior to explain why many people do not exploit cheating opportunities to the full extent, even if they cannot get caught (Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2018; Khalmetski and Sliwka, 2019; Abeler et al., 2019). Our results suggest that people indeed like to be perceived as honest, and that the mere presence of a stranger during the interaction 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; Gächter and Schulz, 2016; Abeler et al., 2019; Cohn et al., 2019).

Second, there is a literature on the evolutionary origins of prosociality, 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 prosocially. These reputational concerns became so deeply ingrained that even the slightest cues of being observed by others can trigger prosocial behavior. Indeed, several studies find that 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).[3] However, the evolutionary legacy hypothesis has also been contested in other studies that 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 social image concerns.

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; Brosig et al., 2003; Bicchieri and Lev-On, 2007), bargaining (e.g., Roth, 1995; Valley et al., 2002), and contract design (Brandts et al., 2016).[4] 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. Abeler et al. (2014) and Conrads and Lotz (2015) also examine honest behavior across different communication channels and do not find significant differences between telephone and computer conditions. However, unlike our experiment, these two studies cannot isolate the role of human presence from the communication mode (i.e., voice vs. text) because both factors were varied simultaneously. Understanding the independent role of human presence in honest behavior is important given the ongoing trend towards automatization. Our study

[3] Relatedly, Hoffman et al. (2015) find that being monitored by a third-party humanized robot reduces cheating to a similar degree as when the observer is human.

[4] A few papers also examine the impact of non-binding promises on trust and trustworthiness (Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006; Vanberg, 2008; Corazzini et al., 2014; Ederer and Stremitzer, 2017).

further departs from the existing literature by showing that dishonest individuals sort themselves into impersonal communication environments.

Experiment 1 – Does machine interaction encourage dishonesty?

Design and procedures

We ran two waves of data collection for the first experiment.[5] The first wave took place in October and November 2013, and the second wave was conducted one year later. The recruitment procedure was the same for both waves (and thus also for each condition). 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). We excluded psychology students as they often participate in experiments that involve deception, and individuals who previously participated in an experiment on lying or cheating. Moreover, we only recruited subjects who had participated at least once in an economic lab experiment to ensure that they trusted our instructions and payment procedure. 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 have a Skype account.[6] We informed them that their personal data would be anonymized for the analysis and treated confidentially, and then obtained their consent to participate in the study. 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 within a wave to avoid differential selection based on initial contact.[7] The experimenter first checked that subjects were in a quiet place, and then told them that they need to get a piece of paper, a pen, and a coin. Subjects then received a link to a short online survey that started with filler questions about life satisfaction and subjective well-being. Subsequently, they 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

[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.

[6] Subjects also had to provide their name, email address, and Skype name as contact details.

[7] In Wave 1, the initial contact was done via Skype call. In Wave 2, subjects were welcomed via Skype chat.

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 E).

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 time (average survey completion time was about 14 minutes). Moreover, since subjects carried out the coin tosses from a remote place, i.e., without being monitored, they could hide behind chance and nobody (including the experimenters) could determine with certainty whether a specific subject misreported their coin tosses. Thus, subjects had a strong financial incentive to cheat without any risk of getting caught. However, reporting a high success rate might be judged as suspicious and undermine a subject's appearance of being an honest person. The coin tossing task and variations of it have been extensively used to study dishonest behavior (see Bucciol and Piovesan, 2011; Houser et al., 2012; Fischbacher and Föllmi-Heusi, 2013; see also Abeler et al., 2019 for a meta-analysis) and the task has been shown to reliably predict rule-violating behavior in natural settings, including violations of prison rules (Cohn et al., 2015; Cingl and Korbel, 2020), misbehavior in school (Cohn and Maréchal, 2018), absenteeism in the workplace (Hanna and Wang, 2017), free riding on public transport (Dai et al., 2018), and adulteration of milk (Kröll and Rustagi, 2016).

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 known (see Houser et al., 2012). Let m be the probability that a subject reports a successful coin flip, conditional on the actual coin toss not being successful (we assume that no one cheats to their disadvantage, i.e., misreports a successful outcome). The probability of reporting a successful outcome p is therefore given by $p = 0.5 \cdot 1 + 0.5 \cdot m = 0.5 \cdot (1 + m)$. If the outcome of a given coin toss is successful, subjects will report a successful outcome with a probability of 1. Conversely, if the coin toss is not successful, subjects will report a successful outcome with a probability of m. Thus, the probability of misreporting the outcome of an unsuccessful coin toss is given by

$$m = 2 \cdot p - 1. \tag{1}$$

Based on the law of large number we can then replace the probability of reporting a successful outcome with the average success rate reported in a given condition to determine the cheating rate in that condition.

For Wave 1, we implemented three treatments that varied, in a between-subjects design, how subjects reported the outcomes of their coin flips (see Table 1). In treatment CALL, subject had to report their outcomes to the experimenter via a Skype call. They were instructed to turn off the video feature in order to keep the degree of anonymity constant across conditions. In treatment FORM, subjects received via Skype a link to a non-interactive online form where they could enter their outcomes. Success rates in these two treatments may therefore differ because of two reasons: (i) the presence or absence of a person during the reporting, and (ii) the type of communication mode (voice vs. text). To this end, we conducted a third condition, treatment CHAT, where subjects were asked to report their outcomes to the experimenter in writing via Skype chat.

In each condition, we informed subjects about the communication channel before they started to toss the coin. Subjects were also told that they will not be asked any questions in response to what they report. A total of n=257 subjects participated in Wave 1 (n=85 in CALL, n=86 in FORM, n=86 in CHAT). The first column of Table A.2 in Online Appendix A presents descriptive statistics of the subjects in Wave 1. They were 24 years old on average and the gender ratio is balanced. Table A.1 describes the sample by treatment and provides randomization checks. The last column in this table indicates that the randomization led to balanced groups, except for the share of medical students (p=0.069, χ^2-test). We therefore control for subjects' fields of study (and other background characteristics) in the regression analysis. The experiment in Wave 1 was run by two female and three male experimenters.

To directly test for the importance of human presence for honest behavior, we conducted a second wave of the experiment. The key condition in Wave 2 is treatment ROBOT where subjects were asked

Table 1. Overview of treatments in Experiment 1

	written communication	oral communication
machine interaction	$\text{FORM}_{1,2}$	ROBOT_{2}
human interaction	CHAT_{1}	$\text{CALL}_{1,2}$

Notes: Subscripts denote whether a treatment was featured in Wave 1 and/or Wave 2.

to use Skype to call a voice response system with pre-recorded voice messages of the experimenters that prompted them to report their outcomes. Since the only difference to CALL is that subjects did not interact with a real person during the reporting, treatment ROBOT allows us isolate the effect of human presence on honest behavior. To ensure a clean comparison, we again conducted treatments CALL and FORM in Wave 2. This allows us to test whether the main findings from Wave 1 replicate. We recruited n=211 new subjects (i.e., subjects who had not participated in Wave 1) for the second wave using the exact same procedure and exclusion criteria as in the first wave (n=67 in CALL, n=75 in FORM, n=69 in ROBOT). Subjects' background characteristics are similar to the first wave with the exception of certain fields of study (there is a higher share of students studying natural sciences and a lower share of students in the "other" category in Wave 2, p=0.008 and 0.018, χ^2-tests; see Table A.2 in Online Appendix A). The last column of Table A.3 confirms that the randomization in Wave 2 was successful, except for the share of students in social sciences (p=0.023, χ^2-test). We address this imbalance by controlling for subjects' fields of study (and other background characteristics) in the regression analysis. The experiment in Wave 2 was run by one female and one male experimenter, both of whom assisted us already in Wave 1. Table 1 presents an overview of the treatments.

There are a few design aspects worth mentioning. First, we designed our experiments so that the reporting of the coin flips was identical across treatments in terms of content. We used the same wording in each condition when we asked subjects to report their outcomes, and they simply had to reply with "Heads" or "Tails" (either orally or in writing) for each coin toss. This permits a *ceteris paribus* comparison of the treatments. Second, we kept subjects' degree of anonymity constant across treatments. In each condition, the experimenters knew a subject's name, email address, and Skype name. Third, we explicitly told all subjects that they will not be asked any questions in response to what they report. Thus, they did not have to worry about having to justify their success rates in any of the treatments.[8] Finally, 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 coin tossing stage

[8] Of course, we can never be entirely sure that subjects believed our instructions. To minimize potential trust issues, we only recruited subjects who had previously participated in economic experiments in the same lab and are thus familiar with the lab's no-deception policy. Yet, it is possible that subjects were concerned about real-time nonverbal reactions from the experimenter (e.g., a cough) that could be interpreted as a sign of suspicion or judgment. While we cannot rule out this possibility, it can be argued that concerns about social disapproval are in line with our proposed mechanism. Individuals who do not care about their social image are also unlikely to respond to nonverbal reactions.

from the reporting stage, and let subjects decide when they wanted to proceed to the reporting stage. This feature minimizes differences in perceived time pressure between treatments.[9]

Results

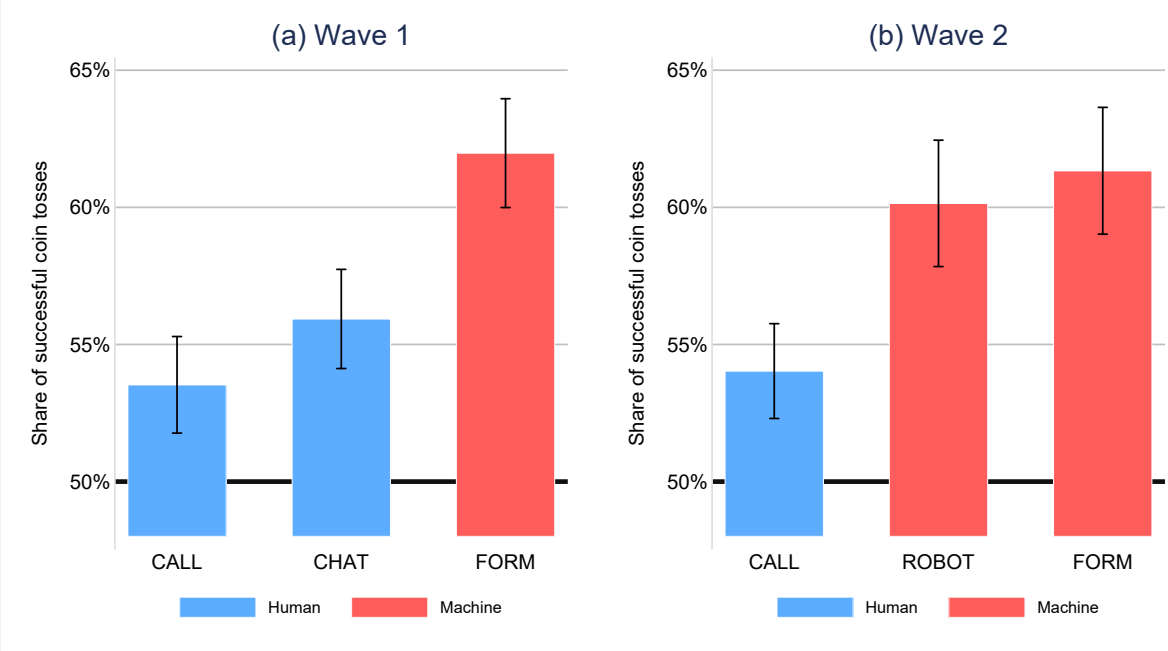
We first report the results from Wave 1. Panel (a) of Figure 1 suggests that subjects were relatively honest when calling the experimenter on Skype to report their outcomes. On average, they reported 53.5% successful coin flips in treatment CALL (95% confidence interval: 50.0–57.0%), which is only slightly higher than the honesty benchmark of 50.0%. According to a simulation analysis with 10,000 coin flipping experiments of the same sample size as in treatment CALL, the probability that subjects reported honestly, but nonetheless achieved a success rate of 53.5% is p=0.0210 (see Online Appendix C for details).[10] Focusing on the entire distribution of outcomes, as shown in Figure 2, we see the largest difference in actual and expected frequencies just above and below the middle outcome (i.e., 5 successful coin flips). In particular, subjects were significantly more likely to report 6 successful outcomes (32.9% vs. 20.5%; p=0.007, binomial test), and were (or tended to be) less likely to report 3 and 4 successful outcomes (2.4% vs. 11.7%, 14.1% vs. 20.5%) compared to the honesty benchmark (p=0.004, and 0.178, binomial tests). This suggests that, while subjects cheated when they reported to a person, they did so only a little bit. Overall, we estimate a cheating rate of 7.0% for CALL (see equation 1).

Cheating was more common when subjects used the online form to report their outcomes. They reported 62.0% successful coin flips, on average, in treatment FORM (95% confidence interval: 58.0–65.9%; p=0.0000, see simulation in Online Appendix C). In particular, we observe a disproportionate share of 10s (7.0% vs. 0.1%), 9s (3.5% vs. 1.0%), 8s (11.6% vs. 4.4%), and 7s (18.6% vs. 11.7%) in FORM compared to the honesty benchmark (p<0.001, 0.052, 0.005, and 0.062, binomial tests; see Figure 2). The cheating rate is 24.0% and, thus, more than three times as large as in CALL. A rank-sum test confirms that the success rates differ significantly between CALL and FORM (p=0.005). As shown in Online Appendix C, the observed difference in success rates between CALL and FORM is very unlikely to have happened by chance. Only six of 10,000 simulated experiments resulted in a similar or larger absolute difference between the two treatments (i.e., p=0.0006).

[9] We empirically address this point in the last paragraph of the "Mechanism" section. The time gap between the coin tossing stage and the reporting stage 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.

[10] We report one-sided p-values for all simulation tests of honest behavior because we assume that people do not cheat to their disadvantage. By contrast, we report two-sided p-values for all simulations of treatment differences.

Figure 1. Information communication technology and cheating

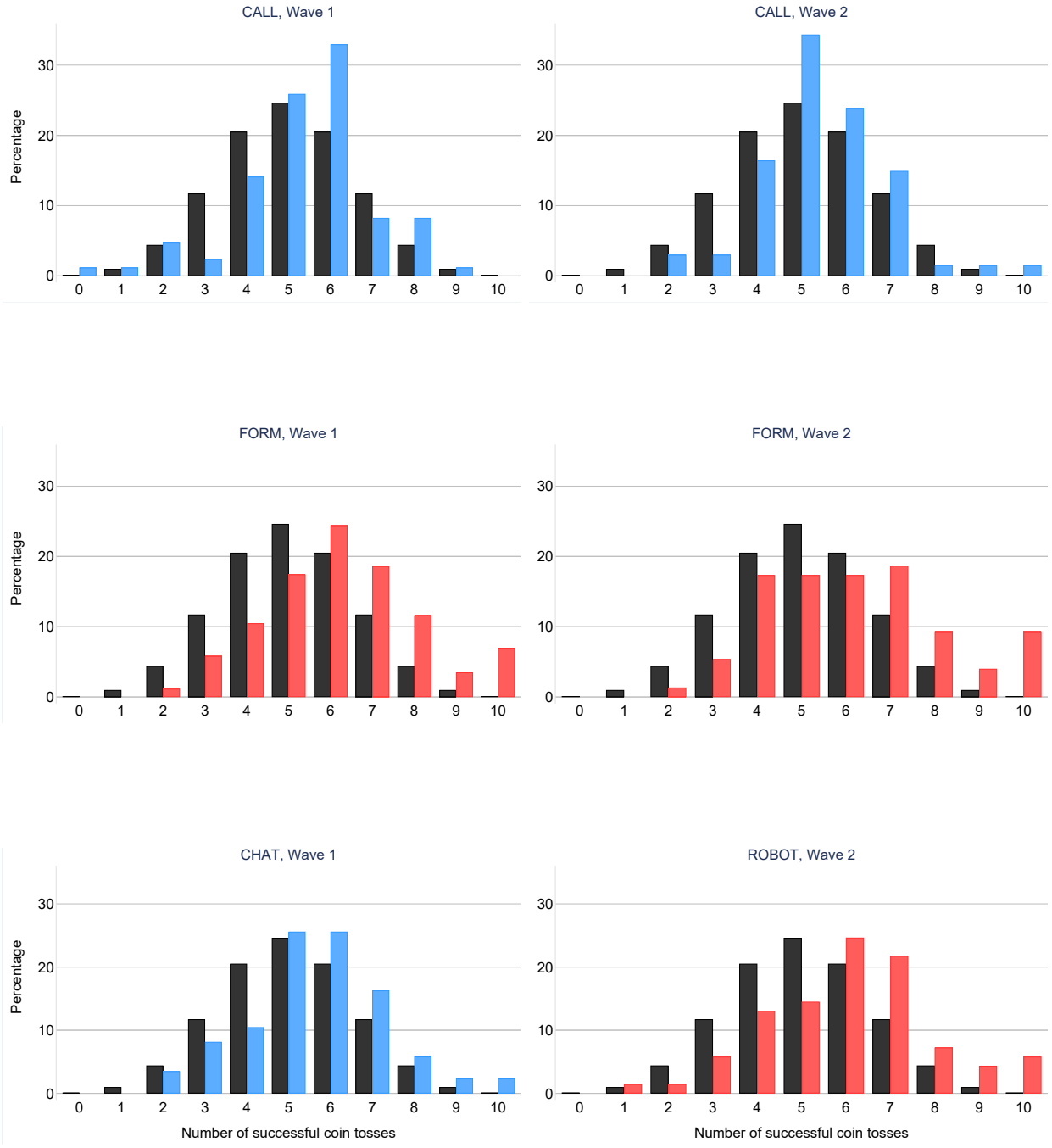


Notes: Percentage of coin tosses reported as successful. Bars indicate standard error of the mean.

We next examine whether the relatively high level of honesty in CALL is due to the communication mode (i.e., voice vs. text). Panel (a) of Figure 1 shows that subjects reported 55.9% successful coin flips, on average, in CHAT (95% confidence interval: 52.3–59.5%; p=0.0001, simulation analysis). Examining the entire distribution in Figure 2, we observe that there are significantly too many 10s (2.3% vs. 0.1%) compared to the honesty benchmark (p=0.003, binomial test). More interestingly, however, we again find a tendency for small acts of cheating when a person was present, as there are too many 6s (25.6% vs. 20.5%) and 7s (16.3% vs. 11.7%), but too few 3s (8.1% vs. 11.7%) and 4s (10.5% vs. 20.5%) relative to what one would expect. However, only the share of 4s is significantly different from the honesty benchmark (p=0.022, binomial test). Overall, we estimate the cheating rate in CHAT at 11.8%. The success rates in CHAT and CALL do not differ significantly (p=0.493, rank-sum test; the corresponding p-value from the simulation is 0.3214). By contrast, the success rate in CHAT is significantly lower than that in FORM despite subjects reporting their outcomes in writing in both conditions (p=0.034, rank-sum test; p=0.0117, simulation analysis). Thus, if anything, the communication mode plays only a minor role in honest behavior.

Next, we look at the results from Wave 2, which introduced treatment ROBOT. Subjects in ROBOT reported their outcomes in the same way as those in CALL, except that there was no actual person present during the reporting. Thus, any difference in cheating between these two conditions can be

Figure 2. Distribution of successful coin tosses by treatments and waves



Notes: Colored bars depict actual observations by treatment; blue=HUMAN (i.e., CALL or CHAT), red=MACHINE (i.e., FORM or ROBOT), and black bars depict the distribution expected under truthful reporting.

attributed to the presence, respectively absence, of a human during the reporting stage. Panel (b) of Figure 1 shows that subjects reported 60.1% successful coin tosses, on average, in ROBOT (95% confidence interval: 55.5–64.7%; p=0.0000, simulation analysis). Figure 2 shows that there are more 10s (5.8% vs. 0.1%), 9s (4.3% vs. 1.0%), and 7s (21.7% vs. 11.7%) compared to what we would expect by chance (p=0.000, 0.030, and 0.015, binomial tests). There are also slightly too many 8s (7.2% vs. 4.4%), but the difference to the honesty benchmark is not significant (p=0.231, binomial test). Overall, the cheating rate in ROBOT amounts to 20.2%.

We find similar results in treatment FORM of Wave 2. Subjects reported 61.3% successful coin tosses, on average (95% confidence interval: 56.7–65.9%; p=0.0000, simulation analysis). This corresponds to a cheating rate of 22.6%. Success rates do not differ significantly between FORM of Wave 2 and ROBOT (p=0.900, rank-sum test; p=0.6534, simulation analysis). By contrast, the success rate in CALL of Wave 2 is only 54.0% and, thus, relatively close to the honesty benchmark (95% confidence interval: 50.6–57.5%; p=0.0197, simulation analysis). Overall, the reported success rates in CALL of Wave 2 translate to a cheating rate of 8.0%, and differ significantly from ROBOT (p=0.025, rank-sum test; p=0.0242, simulation analysis). The fact that the cheating rate in ROBOT is higher than in CALL, but similar to that in FORM, suggests that vocal cues do not activate the same mindset as when people speak to a real person. Rather, the results point to the importance of human presence for honest behavior.

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

$$\Pr\left(y_{it}=1 \mid \boldsymbol{T}_i, \boldsymbol{x}_i, \boldsymbol{z}_j\right) = \Phi\left(\alpha + \beta_1 \text{FORM}_i + \beta_2 \text{ROBOT}_i + \beta_3 \text{CHAT}_i + \boldsymbol{\gamma}'\boldsymbol{x}_i + \boldsymbol{\delta}'\boldsymbol{z}_i\right) \qquad (2)$$

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 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 fields 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. Information communication technology and cheating

	(1)	(2)	(3)	(4)
Dependent variable	$y_{it} = 1$: coin toss reported as successful			
FORM	0.080***	0.080***	0.089***	0.071***
	(0.019)	(0.019)	(0.026)	(0.027)
ROBOT	0.063**	0.069***		0.064**
	(0.025)	(0.025)		(0.027)
CHAT	0.020	0.017	0.025	
	(0.021)	(0.022)	(0.025)	
CALL (base rate)	0.537***	0.537***	0.533***	0.539***
	(0.012)	(0.012)	(0.018)	(0.016)
Controls:				
Subject characteristics	yes	yes	yes	yes
Experimenter FE	no	yes	yes	yes
Wave	1&2	1&2	1	2
Observations	4,680	4,680	2,570	2,110
Subjects	468	468	257	211

Notes: Probit average marginal effects with robust standard errors, corrected for clustering at the individual level, in parentheses. The dependent variable is a dummy indicating whether a subject reported a coin toss as successful (10 observations per subject). The main independent variables are dummies which indicate whether a subject was in treatment FORM, ROBOT, or CHAT (CALL is the level predicted by the model in the reference category). Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 2 presents the regression results. The first two columns are based on the pooled data from both waves (column 1 is without and column 2 is with experimenter fixed effects). The remaining two columns show the results for each wave separately (with experimenter fixed effects). Across both waves, 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; the base rate in CALL is 53.7%).[11]

By contrast, subjects in CHAT were about as likely to report a successful outcome as those in CALL (p=0.347 and 0.434), and significantly less likely (by 6 percentage points) to report a successful outcome compared to those in FORM (p=0.010 and 0.007). The regression results further confirm that subjects cheated to a similar degree in ROBOT and FORM (p=0.540 and 0.683). Comparing columns (1) and (2) indicates that the results remain the same regardless of whether we control for experimenter fixed effects. Moreover, none of the pairwise comparisons of experimenter fixed effects is significant (the smallest p-value is 0.458). Finally, we find the same pattern of results when we analyze the data

[11] If not indicated otherwise, p-values referring to Probit regression outcomes are obtained by the delta-method. Also, in order to provide a meaningful comparison for the reported average marginal effects, base rates are computed as the average value in the baseline category (e.g., CALL in Table 2), as predicted by the model.

by waves, as shown in columns (3) and (4). Together, the results suggest that human presence, rather than the communication mode, affects levels of honesty.

Because we conducted treatments CALL and FORM in each wave, we are able to test how well the results replicate. It turns out that the results are almost identical across waves. In treatment CALL, subjects reported 53.5% and 54.0% successful coin flips in the first and second wave, respectively (p=0.767, rank-sum test). The results from treatment FORM show a similarly high replicability, with 62.0% and 61.3% successful coin flips being reported in the first and second wave, respectively (p=0.710, rank-sum test). We therefore pool the data from the two waves for the remainder of the analysis of Experiment 1.[12]

Mechanism

Why does human interaction promote honest behavior? One possibility is that human presence enhances social image concerns, i.e., individuals' desire to be perceived and judged as honest by others (Abeler et al., 2019; Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2018; Khalmetski and Sliwka, 2019; see also Bursztyn and Jensen, 2017 for a recent review of the social image literature). In our experiment, there was always a human experimenter involved, such as during the welcome stage or for the payments, but only in the human conditions subjects reported their outcomes directly to a person. The experiment thus manipulated subjects' perception of human presence, which we define as the feeling of closeness in terms of socially interacting with another person.[13] Human presence in the reporting stage might have increased subjects' concerns about being judged by others (i.e., the experimenter).

Indeed, an additional survey experiment (N=156) confirms that, compared to human interaction, reporting to a machine significantly reduces both the feeling that another person is present and the desire to leave a good impression on others. We presented our experimental design to respondents on Amazon Mechanical Turk and asked them to take the perspective of a participant in CALL and ROBOT. Subjects then reported how they would perceive the interaction in terms of human presence and how concerned they would be about what the experimenter thinks of them (social image concerns) using multiple

[12]We also investigated whether there are any time trends in the sequence of coin tosses. As shown in Table B.5 in Online Appendix B, we find no evidence for sequence effects. Moreover, we do not find that subjects were initially honest and then later cheated as the correlation between the first five and last five coin tosses across all conditions is positive and significant (Spearman's ρ =0.152, p=0.001).

[13]The manipulation can also be interpreted in terms of social distance. However, given the lack of consensus in the literature on what the term means, we decided to describe the manipulation in terms of human presence. For example, social distance has pre-dominantly been referred to as (i) the degree of subject anonymity or (ii) the extent to which people are similar and share the same social identity. As shown in the next section, our results are difficult to reconcile with these notions of social distance. We thank an anonymous referee for drawing our attention to this.

questionnaire items. The results show that perceptions of human presence are 1.6 standard deviations lower in ROBOT than in CALL, and social image concerns are about 1.5 standard deviations lower ($p<0.001$, t-tests). Moreover, a Blinder-Oaxaca decomposition (Oaxaca, 1973; Blinder, 1973) reveals that the gap in social image concerns can be attributed almost entirely (about 88%) to differences in perceived human presence across treatments (see Online Appendix D for more details).[14]

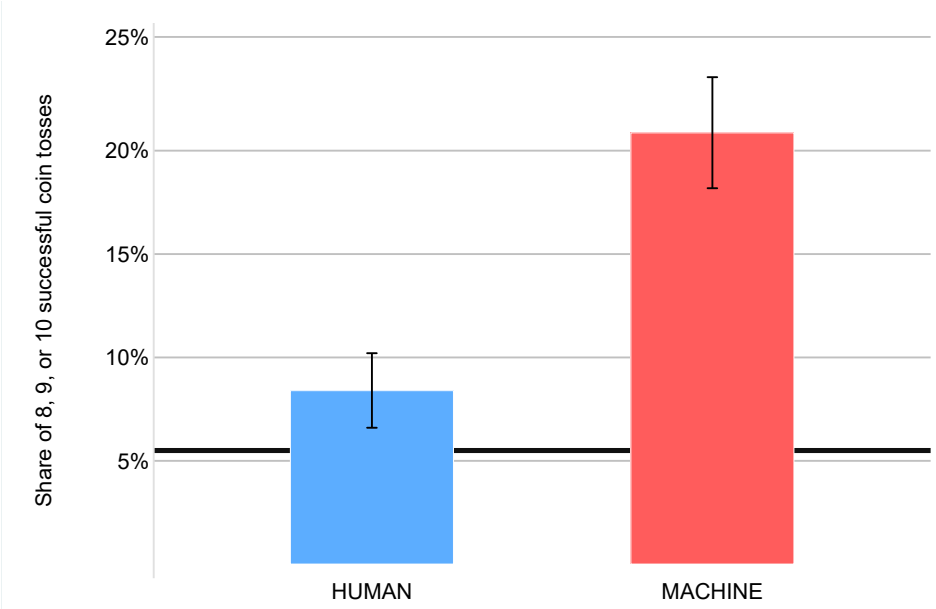
We further provide evidence from our main experiment supporting the notion that social image concerns are driving our treatment effects. If subjects in the machine conditions cared less strongly about their reputation, they should have been more likely to report high, and therefore suspicious, success rates compared to the human conditions. On the other hand, we should not observe a difference in more likely and, thus, more credible outcomes. To test this, we classify 8 or more successful coin flips as suspicious outcomes. The probabilities of 8, 9, and 10 successful coin flips are 4.4%, 1.0%, and 0.1%, respectively. In contrast, we consider 6 and 7 successful coin flips to be more credible because their probabilities are 20.5% and 11.7%, respectively. For ease of exposition, we combine the two human conditions (i.e., CALL and CHAT) as well as the two machine conditions (i.e., FORM and ROBOT) in the analysis.[15]

Figure 3 shows that subjects were more likely to report suspicious outcomes (i.e., 8 or more successes) when they interacted with a machine compared to a human. In the HUMAN conditions, we find that 8.4% of the subjects reported a suspicious outcome (95% confidence interval: 4.9–12.0%). This is just a little bit higher than what we would expect if everyone reported truthfully (5.5%). We additionally conducted simulations, reported in Online Appendix C, showing that given our sample sizes and results, the probability that subjects reported honestly in HUMAN is $p=0.0365$. By contrast, the share of subjects reporting 8 or more successful coin flips in the MACHINE conditions is 20.9% (95% confidence interval: 15.6–26.2%; $p=0.0000$, simulation analysis). Thus, subjects were about two and a half times as likely to report a suspicious success rate in the MACHINE relative to the HUMAN conditions. This suggests that subjects felt more at ease to cheat outright when they reported to a machine compared to a human ($p=0.001$, rank-sum test; $p=0.0000$, simulation analysis).

[14] We also asked subjects at the end of the survey to predict the number of successful coin tosses reported in ROBOT and CALL in the original experiment. We incentivized their predictions by paying a bonus for the most accurate subjects. We find that subjects predicted our treatment effect quite well. On average, they expected participants to report 7.3 successful coin flips in ROBOT and 5.9 in CALL ($p<0.001$, signed-rank test). More importantly, a Blinder-Oaxaca decomposition suggests that about 93% of the difference in predicted coin tossing outcomes between ROBOT and CALL can be explained by differences in the social image index.

[15] For completeness, Table B.1 in Online Appendix B presents the results split by treatments; tables B.2 and B.3 contain the results split by wave.

Figure 3. Suspicious outcomes in HUMAN versus MACHINE interaction



Notes: Percentage of subjects who reported 8 or more successful coin tosses in HUMAN interaction (CALL and CHAT) and MACHINE interaction (FORM and ROBOT). The horizontal line shows the expected rate of 5.5% for 8 or more successful tosses. Bars indicate standard error of the mean.

To examine this pattern in more detail, we estimate Probit models of the following form:

$$\Pr\left(y_i \in Y_s \mid \text{MACHINE}_i, \boldsymbol{x}_i, \boldsymbol{z}_j\right) = \Phi\left(\alpha + \beta_1 \text{MACHINE}_i + \boldsymbol{\gamma}'\boldsymbol{x}_i + \boldsymbol{\delta}'\boldsymbol{z}_i\right) \tag{3}$$

where $\Pr(\cdot)$ denotes the probability that subject i reported a suspicious outcome (e.g., $Y_s = \{8, 9, 10\}$ for our preferred specification). MACHINE_i is an indicator which takes a value of one if the subject reported to a machine (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 in parentheses. As a robustness check, we estimate the same model but shift the threshold for suspicious outcomes by $+/-$ 1 successful coin flips (i.e., the outcome variable in those models is 7 or more and 9 or more successful outcomes). We also perform an analysis of success rates that are more credible, defined as outcomes that are complementary to the suspicious outcomes and that are above 5 (i.e., 6 or 7 successes for our preferred specification, and 6 or 6 to 8 successes for the $+/-$ 1 robustness checks, respectively).

The top panel of Table 3 presents the results for suspicious outcomes, and the bottom panel reports on outcomes that are more credible. Column (1) shows that the share of subjects reporting 8 or more successful coin flips is 15.4 percentage points higher in the MACHINE relative to the HUMAN conditions

Table 3. Suspicious vs. credible outcomes across MACHINE and HUMAN conditions

	(1)	(2)	(3)
	Panel (a): Suspicious over-reporting		
Dependent variable:	$y_i \in \{8,9,10\}$	$y_i \in \{7,8,9,10\}$	$y_i \in \{9,10\}$
MACHINE	0.154***	0.207***	0.093***
	(0.032)	(0.042)	(0.024)
Base rate	0.074***	0.207***	0.027***
	(0.016)	(0.026)	(0.010)
Expected rate	0.055	0.172	0.011
	Panel (b): Credible over-reporting		
Dependent variable:	$y_i \in \{6,7\}$	$y_i \in \{6\}$	$y_i \in \{6,7,8\}$
MACHINE	-0.008	-0.067*	0.055
	(0.046)	(0.040)	(0.047)
Base rate	0.416***	0.283***	0.460***
	(0.032)	(0.029)	(0.032)
Expected rate	0.321	0.205	0.367
Controls:			
Subject characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	1&2	1&2	1&2
Observations	468	468	468

Notes: Average marginal effects of a Probit regression with robust standard errors in parentheses. The dependent variable is a dummy which indicates whether y_i, the number of successful coin tosses reported by a subject, is within the respective sets. The main independent variable MACHINE is a dummy which indicates whether a subject reported to a machine (FORM or ROBOT). The two treatments with human interaction (CALL and CHAT) serve as the reference category. "Base rate" refers to the proportion of positive outcomes for the dependent variable which the regression model predicts for the reference category. "Expected rate" refers to the outcome for the dependent variable that is expected under truthful reporting. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Data from Wave 1 and Wave 2 are pooled. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

(p<0.001). Columns (2) and (3) show that the results are robust to variations in the threshold. The difference in reporting suspicious outcomes is 20.7 percentage points if we set the threshold at 7 or more successful coin flips (p<0.001; see column 2), and it is 9.3 percentage points if the threshold is set at 9 or more successful coin flips (p<0.001; see column 3).

By contrast, we do not observe such a pattern for outcomes that are more likely and, thus, more credible. Column (1) in the bottom panel of Table 3 shows that, while subjects reported 6 or 7 successful coin flips more often than predicted by chance (41.6% instead of 32.2%; 95% confidence interval: 35.3–47.9%), they did so to a similar extent regardless of whether they reported to a machine or a person (p=0.862). The results do not meaningfully change when we use alternative thresholds for credible

outcomes, as shown in columns (2) and (3) (p=0.092 and 0.241). Together, these results suggest that subjects felt more comfortable to report suspicious success rates when they interacted with a machine compared to a human—a finding that is consistent with the notion that human presence enhances people's desire to maintain a positive social image.

Alternative explanations

We now explore several alternative explanations for why subjects cheated less when they interacted with a human rather than a machine. First, while the detection probability was effectively zero in all conditions, it is nonetheless conceivable that some subjects erroneously thought that they could get caught cheating and, consequently, that they would not get paid. If true, we should observe that our results vary based on subjects' risk attitudes. Specifically, we should see that more risk-averse subjects (i) are generally less likely to cheat, and (ii) that they react more strongly to the presence of the experimenter as they are presumably more concerned about getting caught when interacting with a person.

To examine this, we elicited subjects' risk attitudes using an experimentally validated survey question developed by Dohmen et al. (2011) (see also Falk et al., 2018; Schürmann et al., 2018, for additional validations of this questionnaire measure). 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 reverse-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 similar to (3) but with two changes: (i) the dependent variable is whether subject i reported coin toss t as successful ($y_{it} = 1$), and (ii) we add our measure of risk aversion along with its interaction with the MACHINE dummy.[16] The interaction term allows us to test whether more risk-averse subjects are more sensitive to the presence of a person (i.e., the experimenter).

Table 4 presents the results in three steps. Column (1), which does not control for subjects' risk aversion, shows that subjects were about 7.2 percentage points more likely to report a successful outcome when the reporting was made to a machine rather than a person (p<0.001). Column (2) indicates that across conditions, subjects' risk aversion does not significantly predict their likelihood of

[16] Table B.4 in Online Appendix B suggests that our results do not change if we interact risk aversion with individual treatment indicators instead of the MACHINE dummy.

Table 4. Risk aversion and cheating

	(1)	(2)	(3)
Dependent variable	$y_{it} = 1$: coin toss reported as successful		
MACHINE	0.072***	0.072***	0.074***
	(0.016)	(0.016)	(0.016)
Risk aversion		-0.003	-0.018
		(0.009)	(0.013)
Risk aversion × MACHINE			0.031
			(0.021)
Controls			
Subject Characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	1&2	1&2	1&2
Observations	4,680	4,680	4,680
Subjects	468	468	468

Notes: Probit average marginal effects with robust standard errors, corrected for clustering at the individual level, 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 (FORM or ROBOT). The two treatments with human interaction (CALL and CHAT) serve as the reference category. Risk aversion is based on subjects' 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 recoded this measure such that larger values indicate higher risk aversion and then normalized it so that the variable "Risk aversion" has a mean of zero and a standard deviation of one. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Data from Wave 1 and Wave 2 are pooled. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

reporting a successful outcome (p=0.756). This suggests subjects were not worried about punishment for misreporting their outcomes. Moreover, controlling for risk aversion does not change the coefficient of MACHINE.

Column (3) reports the results when we additionally include the interaction between risk aversion and MACHINE. 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 subjects felt significantly more comfortable cheating in the machine conditions. However, we find that the coefficient of risk aversion remains small and insignificant in either the MACHINE or HUMAN conditions (p=0.172 and p=0.339). Also, the difference between the two coefficients is not significant (p=0.138). Overall, the results do not support the conjecture that punishment concerns drive the difference between the MACHINE and HUMAN conditions.

Second, we investigate whether our findings can be explained by differences in time pressure across conditions. Perhaps subjects felt more pressure when they reported their outcomes to the experimenter and, consequently, cheated less (see Capraro, 2017; Lohse et al., 2018). Yet, our design minimized this possibility in two ways. We informed subjects that they will not be asked any follow-up questions about what they report. Moreover, 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 directly analyze subjects' perceived time pressure using responses to the question "To what extent did you feel under time pressure when reporting the outcomes of your coin tosses?". Answers were elicited on a 7-point Likert scale ranging from "not at all" (=0) to "very much" (=6). We find that 78.4% 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).[17] Thus, it is unlikely that subjects cheated less in the HUMAN conditions due to higher time pressure.

Finally, a large literature suggests that increased social distance has a negative impact on prosocial and cooperative behavior.[18] This literature encompasses two distinct conceptualizations of social distance: (i) the degree of subject anonymity (e.g., Hoffman et al., 1996; Bohnet and Frey, 1999; Andreoni and Petrie, 2004; Charness and Gneezy, 2008)[19] and (ii) the extent to which people are similar and share the same social identity (e.g., Akerlof, 1997; **?**; Buchan et al., 2006; Charness et al., 2007; Goeree et al., 2010). However, both notions of social distance are distinct from what we consider as human presence (i.e., the feeling of closeness in terms of socially interacting with another person). As described in more detail below, we find no indication that our results are driven by social distance, neither in terms of anonymity nor in terms of similarity.

Subject anonymity: The experimenters knew subjects' names, email addresses, and their Skype names in all treatments. Thus, since there is no variation in subject anonymity across conditions, it cannot explain why subjects were more likely to cheat when interacting with a machine.

[17] The corresponding p-values for the individual treatment comparisons are: p=0.861 for FORM vs. CALL, p=0.148 for FORM vs. CHAT, p=0.660 for ROBOT vs. CALL, and p=0.349 for ROBOT vs. CHAT (rank-sum tests).

[18] Two recent studies also examined the influence of social distance on cheating (see Hermann and Ostermaier, 2018; Heymann and Rey Biel, 2018).

[19] It is not without controversy whether subject anonymity is an appropriate proxy for social distance. For example, Dufwenberg and Muren (2006) argue on p. 46 that "[...] *it is problematic to organize experimental data in terms of social distance if this notion is taken to vary one-to-one with anonymity. As anonymity changes other things may change alongside so that confounding factors may inadvertently be introduced.*"

Similarity: It is conceivable that hearing the experimenter's voice in treatment CALL revealed additional information about their social identity. For example, subjects might have learned that they share the same gender or cultural background as the experimenter.[20] However, since in all treatments of Wave 1 the experimenters first contacted the subjects by calling them on Skype, subjects had exactly the same information about the experimenters' background in each condition. Only in Wave 2, where subjects were welcomed via text messages in Skype chat, subjects might have had different information about the experimenters' background and, thus, also different perceptions of their similarity or social distance. We examine this possibility in Table B.6 in the Online Appendix, where we restrict the sample to treatments CALL and FORM of Wave 2. Specifically, we interact treatment CALL with three dummy variables for whether a subject and the experimenter had the same gender, native language, or both. If social distance drives our treatment effects, we should observe that subjects with lower social distance to the experimenter (i.e., those who had more in common with the experimenter) reacted more strongly to the presence of the experimenter in CALL. However, we find that none of the corresponding interaction effects are significant (p=0.665, p=0.102, and p=0.579), and the coefficients also do not have the expected sign. We further examine whether, in general, similarities between a subject and the experimenter influenced cheating using data from all treatments where subjects could hear the experimenter's voice (either during the welcome stage or the reporting stage). Table B.7 in the Online Appendix shows that none of the social distance proxies (same gender, same native language, or both same gender and same language) are significantly related to subjects' reporting behavior (p=0.511, p=0.285, and p=0.314).

Experiment 2 – Do dishonest people prefer machine interaction?

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 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 a machine.

[20] The experimenters were all Swiss-German and, thus, speak German with a distinctive accent. Swiss-German is also the native language of 66.7% of our subjects.

Design and procedures

We recruited a new sample of subjects (i.e., subjects who had not participated in Experiment 1) 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 at the time when they signed up for the study. 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.[21] 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 paid based on their responses in that part. We explained the payment procedure to the participants and assured them, before obtaining their informed consent, that their data will be anonymized for the analysis and treated confidentially.

For Part A, subjects 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 Experiment 1. And 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 subjects' earnings from the initial coin tossing task as a proxy for their 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 a Skype call (without video) or they could use the online form. The two options were presented in randomized order.

Subjects who were selected to participate 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

[21] We limited the number of participants for Part B because our main focus are choices of the communication channel in Part A.

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 Experiment 2 (88 subjects in Part B of Experiment 2, respectively). They were 23 years old, on average, and 47.4% were male (see Table A.4 in Online Appendix A). We employed one experimenter for Part B.

Results

The results from Part A reveal that subjects cheated to some extent. They reported, on average, 58.4% successful coin flips (95% confidence interval: 56.6—60.3%). The simulation analysis reported in Online Appendix C suggest that this result is inconsistent with completely truthful reporting (p=0.0000).

For Part B we find that 50.5% of the subjects chose to call the experimenter on Skype to report their outcomes of the second coin tossing task, and 49.5% of them chose the online form. The binned scatter plot in Figure 4 (following the procedure of Chetty et al., 2014) shows that subjects who reported a higher number of successful coin flips in Part A of the experiment (i.e., those who presumably cheated) were also more likely to choose the online form for Part B.

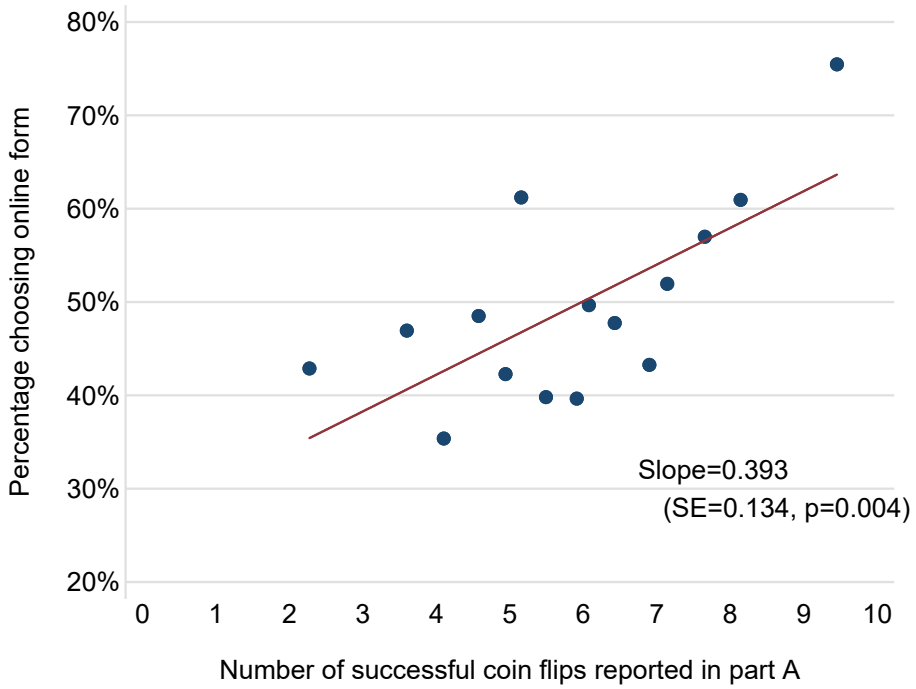
We corroborate these results by estimating a Probit model of the following form:

$$\Pr\left(c_i = \text{FORM} \mid y_i^A, \boldsymbol{x}_i\right) = \Phi\left(\alpha + \beta_1 y_i^A + \boldsymbol{\gamma}'\boldsymbol{x}_i\right) \tag{4}$$

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 3.8 percentage points more likely to select the online form for Part B (p=0.005). The results remain the same when we control for subjects' background characteristics, as shown in column 2 (p=0.003). By contrast, none of subjects' background characteristics (i.e., age, gender, nationality, and fields of study) predicts their choices of the reporting channel significantly at the 5% level. In column (3) we replaced the number of successful coin flips with a variable indicating how many successful coin tosses subjects reported in excess of 5 (the variable has a value of zero if the outcome is five or below). The results suggest that for every successful coin flip above 5 in Part A, the likelihood that a subject prefers to report through the online form in Part B increases by 5.9 percentage points (p=0.001). In sum, when given the choice, alleged cheaters prefer to avoid human interaction.

Figure 4. Screening for dishonest people



Notes: 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 the online form to report outcomes in Part B. 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. 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.

Table 5. Selection into machine reporting

	(1)	(2)	(3)
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)	
Successful coin tosses in Part A (>5)			0.059***
			(0.017)
Controls:			
Age (years)		0.011	0.012
		(0.007)	(0.007)
Male subject		-0.084	-0.082
		(0.053)	(0.053)
Swiss nationality		0.036	0.041
		(0.058)	(0.058)
Field of study: Law		-0.005	-0.018
		(0.109)	(0.107)
Field of study: Economics/Business		-0.094	-0.092
		(0.097)	(0.098)
Field of study: Medicine		0.148	0.146
		(0.097)	(0.097)
Field of study: Social Sciences		0.212*	0.219*
		(0.120)	(0.119)
Field of study: Natural Sciences		0.007	0.006
		(0.060)	(0.059)
Observations	380	380	380

Notes: Probit average marginal effects with robust standard errors 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. In columns (1) and (2), the main independent variable is the number of successful coin tosses reported in Part A. In column (3), this variable indicates how many coin tosses are in excess of 5 (i.e., the variable is zero for outcomes of 5 or less). Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Note that dishonest people might avoid human interactions for reasons other than their increased propensity to cheat. For example, it could be that dishonest people are more likely to be introverted and that they prefer to avoid human interactions because of their introversion rather than their disposition to lie. Thus, failing to control for such personality traits may lead to a biased interpretation of the results. Yet, even if people avoid human interactions because of some unobserved personality traits, offering the choice between human and machine interaction could still be an effective tool to screen for dishonest people to the extent that the personality traits correlate with people's tendency to cheat.

We examine the possibility of omitted variable bias using a method proposed by Oster (2019). The approach is based on the assumption that the relationship between the independent variable of interest (in our case, it is the number of successful outcomes reported in Part A) and unobservables can be recovered from the relationship between the variable of interest and observables. More specifically, the extent of omitted variables bias can be assessed by estimating the sensitivity of the coefficient of interest to the inclusion of observed controls, relative to the change in the R-squared when the controls are included. Applying this method, we find that the relationship between unobservables and the outcome (i.e., subjects' choice of communication channel) would need to be about 35 times stronger (and exhibit the opposite sign) than the relationship between observables and the outcome in order to drive our estimated coefficient of interest from 5.9 percentage points down to zero.[22]

While cheating behavior in Part B was not the focus of Experiment 2, we nonetheless report the results for completeness (see Figure B.1 in the Online Appendix for the full distributions of reported outcomes in Part A and B). Note that the sample size in Part B is substantially smaller than in Part A because only a quarter of subjects were invited at random for Part B and 12% of those invited did not show up. The following results should therefore be interpreted with caution. We find that subjects who chose to call the experimenter in Part B reported 59.7% successful coin flips on average (95% confidence interval: 55.5—64.0%; p=0.0001, simulation in Online Appendix), and those who chose the online form reported 65.9% successes (95% confidence interval: 60.6—71.1%; p=0.0000, simulation analysis).[23] The difference in reporting behavior between subjects who chose CALL and those who

[22]Specifically, we estimate a linear probability model using the same specification as in Table 5. To account for measurement error, we follow the recommendation of Oster (2019) and multiply the obtained R-squared obtained from that regression with 1.3 to get an estimate for the R-squared from a (hypothetical) regression that includes both the observed and unobserved control variables.

[23]The Spearman's ρ for the correlation of success rates between Part A and B is 0.388 ($p<0.001$), providing additional evidence that subjects cheated. Success rates in Part B were higher than those we find in Experiment 1 and Part A of Experiment 2. This could be due to disproportionate attrition of honest subjects (as they will earn less in expectation)

chose FORM is marginally significant based on non-parametric tests (p=0.085, rank-sum test; p=0.0726 simulation analysis). However, when we use regression to control for background characteristics we find that subjects in Part B were significantly more likely (by 8.4 percentage points) to report a successful outcome when they chose to report via the online form as opposed to calling the experimenter on Skype (p=0.012; see Table B.8).

Conclusion

The digital age has radically changed the way we communicate and interact with each other. For example, 50 years ago we walked over to the local branch of the insurance company to report a stolen bicycle, 20 years ago we called a representative of the insurance company, and today we can just fill out an online form or chat with a bot. Are we more likely to misrepresent information when we submit 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 advantage. Our experimental paradigm for measuring dishonest behavior is a coin tossing task in which subjects are asked to privately flip a coin multiple times, report the outcomes of their 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 flip outcomes and found that they cheated substantially less when they interacted with a person rather than a machine.[24] Human presence appears to be essential for honest reporting because adding human features (i.e., a human voice) to a machine does not encourage more honest behavior.[25] Further analysis and an additional survey experiment suggests that human presence enhances social image concerns, i.e., individuals' desire to maintain an honest appearance even if they will never meet the other person again. Our findings suggest that interaction with humans is key to reducing fraudulent behavior, which is relevant to any organization that relies on customers' or employees' willingness to behave honestly. 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 presence in mitigating

or experience effects (see also Fischbacher and Föllmi-Heusi, 2013, for evidence that repeated participation in cheating experiments increases cheating).

[24] Using observational data, Laudenbach et al. (2018) provide recent evidence corroborating the importance of human interaction for honest behavior. They find that bank customers who receive a personal call from their bank agent are less likely to default on their loans than those who only get a letter.

[25] An interesting avenue for future research is to explore other ways of humanizing robots, such as language style and appearance (Araujo, 2018; Siebenaler et al., 2019), and test how these human features affect cheating in human-robot interaction. We thank an anonymous reviewer for pointing this out.

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 are also more likely to avoid communication channels that require them to interact with a real person. We indeed find that subjects who are more likely to cheat prefer to report their coin flip outcomes to a machine. This finding raises the possibility for organizations to screen for customers or employees with an increased propensity to engage in fraudulent behavior. For example, firms could interact with their customers via multiple communication channels that differ by whether customers interact with a real agent, and then focus on suspicious cases where the customers choose to avoid human interaction. Thus, offering the option of machine-based customer service has the potential to reduce personnel and antifraud investigation costs.[26] However, it may also introduce the risk of attracting new customers who are particularly dishonest, which can ultimately hurt companies that provide such multi-platform customer service.[27] Future research is needed to determine the long-run consequences of impersonal customer and employee interaction.

[26] Of course, the effectiveness of such a screening approach will depend on how predictive people's choices of communication channels are for their tendency to cheat and this may vary across contexts.

[27] The possibility of self-selection of individuals into different institutional environments can have important general equilibrium effects (see, e.g., Gürerk et al., 2006 for an application in the context of public goods, and Houdek, 2017 for a discussion specific to dishonest behavior).

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 (2019). Preferences for Truth-Telling. *Econometrica 87*(4), 1115–1153.

Acemoglu, D. and P. Restrepo (2017). Robots and Jobs: Evidence from US Labor Markets. *NBER Working Paper 23285*.

Akerlof, G. a. (1997). Social Distance and Social Decisions. *Econometrica 65*(5), 1005–1027.

Andreoni, J. and R. Petrie (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics 88*(7-8), 1605–1623.

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior 85*, 183–189.

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.

Aron, A., E. Aron, and D. Smollan (1992). Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness. *Journal of Personality and Social Psychology 63*(4), 596–612.

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.

Blinder, A. S. (1973). Wage Discrimination : Reduced Form and Structural Estimates. *The Journal of Human Resources 8*(4), 436–455.

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).

Brosig, J., A. Ockenfels, and J. Weinmann (2003). The effect of communication media on cooperation. *German Economic Review 4*(2), 217–241.

Bucciol, A. and M. Piovesan (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology 32*(1), 73–78.

Buchan, N., E. Johnson, and R. Croson (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization 60*, 373–398.

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. *Quarterly Journal of Economics 133*(3), 1561–1595.

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.

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 & 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.

Cingl, L. and V. Korbel (2020). External validity of a laboratory measure of cheating: Evidence from Czech juvenile detention centers. *Economics Letters 191*, 109094.

Cohn, A. and M. A. Maréchal (2018). Laboratory Measure of Cheating Predicts Misbehavior at School. *Economic Journal 128*(65), 2743–2754.

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.

Cohn, A., M. A. Maréchal, D. Tannenbaum, and C. L. Zünd (2019). Civic honesty around the globe. *Science 365*(6448), 70–73.

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 (2018). Cheating in the Lab Predicts Fraud in the Field. An Experiment in Public Transportations. *Management Science 64*(3), 1081–1100.

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.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

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

Dufwenberg, M. and A. Muren (2006). Generosity, anonymity, gender. *Journal of Economic Behavior and Organization 61*(1), 42–49.

Ederer, F. and A. Stremitzer (2017). Promises and expectations. *Games and Economic Behavior 106*, 161–178.

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.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global Evidence on Economic Preferences. *Quarterly Journal of Economics 133*(4), 1645–1692.

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.

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in Disguise-an Experimental Study on Cheating. *Journal of the European Economic Association 11*(3), 525–547.

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 (2018). Lying Aversion and the Size of the Lie. *American Economic Review 108*(2), 419–453.

Goeree, B. J. K., M. A. Mcconnell, T. Mitchell, T. Tromp, and L. Yariv (2010). The 1/d Law of Giving† By. *American Economic Journal: Microeconomics 2*(1), 183–203.

Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2006). The Competitive Advantage of Sanctioning Institutions. *Science 312*(5770), 108–111.

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*.

Hancock, J. T. and J. Guillory (2015). Deception with Technology. In S. S. Sunder (Ed.), *The Handbook of the Psychology of Communication Technology*, Chapter 12, pp. 270–289. Wiley & Sons.

Hancock, J. T., J. Thom-Santelli, and T. Ritchie (2004). Deception and Design: The Impact of Communication Technology on Lying Behavior. In *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 129–134.

Hanna, R. and S.-Y. Wang (2017). Dishonesty and selection into public service. *American Economic Journal: Economic Policy 9*(3), 262–290.

Hermann, D. and A. Ostermaier (2018). Be Close to Me and I Will Be Honest. How Social Distance Influences Honesty. *CEGE Discussion Paper* (340).

Heymann, D. C. and P. Rey Biel (2018). Social Distance in Deceptive Behavior: Experimental Evidence from Malawi. *mimeo*.

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.

Hoffman, G., J. Forlizzi, S. Ayal, A. Steinfeld, J. Antanitis, G. Hochman, E. Hochendoner, and J. Finkenaur (2015). Robot Presence and Human Honesty: Experimental Evidence. In *10th ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*, pp. 181–188. ACM.

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.

Houdek, P. (2017). A perspective on research on dishonesty: Limited external validity due to the lack of possibility of self-selection in experimental designs. *Frontiers in Psychology 8*(1566), 1–6.

Houser, D., S. Vetter, and J. Winter (2012). Fairness and cheating. *European Economic Review 56*(8), 1645–1655.

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 (2019). Disguising Lies - Image Concerns and Partial Lying in Cheating Games. *American Economic Journal: Microeconomics forthcomin*.

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*.

Laudenbach, C., J. Pirschel, and S. Siegel (2018). Personal Communication in a Fintech World: Evidence from Loan Payments. *mimeo*.

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.

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.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review 14*(3), 693–709.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business and Economic Statistics 37*(2), 187–204.

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.

Schürmann, O., S. Andraszewicz, and J. Rieskamp (2018). The Importance of Losses when Eliciting Risk Preferences. *mimeo*.

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.

Siebenaler, S., A. Szymkowiak, P. Robertson, G. I. Johnson, J. Law, and K. Fee (2019). Honesty, Social Presence and Self-Service in Retail. *Interacting with Computers 31*(2), 154–166.

Toma, C., J. Bonus, and L. Van Swol (2019). Lying Online: Examining the Production, Detection, and Popular Beliefs Surrounding Interpersonal Deception in Technologically-Mediated Environments. In *The Palgrave Handbook of Deceptive Communication.*, Chapter 31, pp. 583–601. Palgrave Macmillan.

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.

For Online Publication

Honesty in the Digital Age

Online Appendix

Alain Cohn, Tobias Gesche, and Michel Maréchal[1]

Contents:

[1]Cohn: adcohn@umich.edu, School of Information, University of Michigan; Gesche: tgesche@ethz.ch, Center for Law & Economics, ETH Zurich; Maréchal: michel.marechal@econ.uzh.ch, Department of Economics, University of Zurich

Appendix A: Descriptive statistics and randomization checks

Table A.1. Descriptive statistics and randomization checks for Wave 1 of Experiment 1

	CALL	CHAT	FORM	p-value
Age (years)	24.600	23.163	23.047	0.195
	(6.461)	(2.656)	(4.351)	
Male subject	0.518	0.534	0.453	0.531
	(0.503)	(0.501)	(0.508)	
Swiss nationality	0.706	0.779	0.755	0.532
	(0.458)	(0.417)	(0.432)	
Field of study: Law	0.059	0.058	0.035	0.718
	(0.237)	(0.235)	(0.185)	
Field of study: Economics/Business	0.047	0.116	0.070	0.226
	(0.213)	(0.322)	(0.256)	
Field of study: Medicine	0.047	0.034	0.116	0.069
	(0.213)	(0.185)	(0.322)	
Field of study: Social Sciences	0.153	0.105	0.186	0.319
	(0.362)	(0.308)	(0.391)	
Field of study: Natural Sciences	0.447	0.384	0.326	0.264
	(0.500)	(0.489)	(0.471)	
Field of study: Other	0.247	0.302	0.267	0.714
	(0.434)	(0.462)	(0.445)	
Observations	85	86	86	

Notes: This table reports means and standard deviations (in parentheses) of subjects' age (in years), gender (1=male), Swiss citizens (1=yes), and fields of study. The last column contains p-values for the null hypothesis of perfect randomization (Kruskal-Wallis test for age and χ^2-tests for all other variables).

Table A.2. Descriptive statistics across for each wave of Experiment 1

	Wave 1	Wave 2	p-value
Age (years)	23.599	24.123	0.711
	(4.778)	(6.129)	
Male subject	0.502	0.474	0.546
	(0.501)	(0.501)	
Swiss nationality	0.747	0.754	0.872
	(0.436)	(0.432)	
Field of study: Law	0.051	0.047	0.874
	(0.220)	(0.213)	
Field of study: Economics/Business	0.078	0.104	0.319
	(0.268)	(0.306)	
Field of study: Medicine	0.066	0.057	0.679
	(0.249)	(0.232)	
Field of study: Social Sciences	0.148	0.104	0.160
	(0.356)	(0.306)	
Field of study: Natural Sciences	0.385	0.507	0.008
	(0.488)	(0.501)	
Field of study: Other	0.272	0.180	0.018
	(0.446)	(0.385)	
Observations	257	211	468

Notes: This table reports means and standard deviations (in parentheses) of subjects' age (in years), gender (1=male), Swiss citizens (1=yes), and fields of study. The last column contains p-values for the null hypothesis of perfect randomization (Kruskal-Wallis test for age and χ^2-tests for all other variables).

Table A.3. Descriptive statistics and randomization checks for Wave 2 of Experiment 1

	CALL	ROBOT	FORM	p-value
Age (years)	23.478	24.478	24.373	0.409
	(4.974)	(5.994)	(7.139)	
Male subject	0.507	0.464	0.453	0.795
	(0.503)	(0.502)	(0.501)	
Swiss nationality	0.761	0.724	0.773	0.783
	(0.429)	(0.450)	(0.421)	
Field of study: Law	0.060	0.043	0.040	0.844
	(0.239)	(0.205)	(0.197)	
Field of study: Economics/Business	0.075	0.072	0.160	0.144
	(0.265)	(0.261)	(0.369)	
Field of study: Medicine	0.030	0.043	0.093	0.223
	(0.171)	(0.205)	(0.293)	
Field of study: Social Sciences	0.149	0.145	0.027	0.023
	(0.359)	(0.354)	(0.162)	
Field of study: Natural Sciences	0.552	0.522	0.453	0.479
	(0.501)	(0.503)	(0.501)	
Field of study: Other	0.134	0.174	0.227	0.355
	(0.344)	(0.382)	(0.421)	
Observations	67	69	75	

Notes: This table reports means and standard deviations (in parentheses) of subjects' age (in years), gender (1=male), Swiss citizens (1=yes), and fields of study. The last column contains p-values for the null hypothesis of perfect randomization (Kruskal-Wallis test for age and χ^2-tests for all other variables).

Table A.4. Descriptive statistics for Experiment 2

Age (years)	23.089
	(3.725)
Male subject	0.474
	(0.500)
Swiss nationality	0.742
	(0.438)
Field of study: Law	0.066
	(0.248)
Field of study: Economics/Business	0.084
	(0.278)
Field of study: Medicine	0.082
	(0.274)
Field of study: Social Sciences	0.047
	(0.213)
Field of study: Natural Sciences	0.350
	(0.478)
Field of study: Other	0.371
	(0.484)
Observations	380

Notes: This table reports means and standard deviations (in parentheses) of subjects' age (in years), gender (1=male), Swiss citizens (1=yes), and fields of study.

Appendix B: Robustness checks and additional tables and figures

Table B.1. Suspicious vs. credible outcomes across treatments

	(1)	(2)	(3)
	Panel (a): Suspicious over-reporting		
Dependent variable:	$y_i \in \{8,9,10\}$	$y_i \in \{7,8,9,10\}$	$y_i \in \{9,10\}$
FORM	0.177***	0.234***	0.131***
	(0.047)	(0.053)	(0.042)
ROBOT	0.198***	0.256***	0.169**
	(0.070)	(0.070)	(0.071)
CHAT	0.015	0.073	0.042
	(0.052)	(0.065)	(0.050)
Base rate	0.070***	0.185***	0.019*
	(0.019)	(0.031)	(0.010)
Expected rate	0.055	0.172	0.011
	Panel (b): Credible over-reporting		
Dependent variable:	$y_i \in \{6,7\}$	$y_i \in \{6\}$	$y_i \in \{6,7,8\}$
FORM	-0.008	-0.078*	0.052
	(0.055)	(0.045)	(0.056)
ROBOT	0.045	-0.050	0.096
	(0.074)	(0.057)	(0.073)
CHAT	0.045	-0.016	0.028
	(0.068)	(0.056)	(0.068)
Base rate	0.400***	0.289***	0.450***
	(0.039)	(0.036)	(0.040)
Expected rate	0.322	0.205	0.367
Controls:			
Subject characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	1&2	1&2	1&2
Observations	468	468	468

Notes: Probit average marginal effects with robust standard errors in parentheses. The dependent variable is a dummy which indicates whether y_i, the number of successful coin tosses reported by a subject, is within the respective sets. The main independent variable are dummies which indicate whether a subject was in either treatments FORM, ROBOT, CHAT (CALL is the reference category). "Base rate" refers to the proportion of positive outcomes for the dependent variable which the regression model predicts for the reference category. "Expected rate" refers to the outcome for the dependent variable that is expected under truthful reporting. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Data from Wave 1 and Wave 2 are pooled. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B.2. Suspicious vs. credible outcomes across MACHINE and HUMAN treatments in Wave 1

	(1)	(2)	(3)
	Panel (a): Suspicious over-reporting		
Dependent variable:	$y_i \in \{8,9,10\}$	$y_i \in \{7,8,9,10\}$	$y_i \in \{9,10\}$
MACHINE	0.127***	0.184***	0.085**
	(0.047)	(0.062)	(0.035)
Base rate	0.098***	0.223***	0.027**
	(0.022)	(0.032)	(0.011)
Expected rate	0.055	0.172	0.011
	Panel (b): Credible over-reporting		
Dependent variable:	$y_i \in \{6,7\}$	$y_i \in \{6\}$	$y_i \in \{6,7,8\}$
MACHINE	0.002	-0.056	0.056
	(0.064)	(0.055)	(0.065)
Base rate	0.420***	0.296***	0.487***
	(0.037)	(0.033)	(0.037)
Expected rate	0.322	0.205	0.367
Controls:			
Subject characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	1	1	1
Observations	257	257	257

Notes: Probit average marginal effects with robust standard errors in parentheses. The dependent variable is a dummy which indicates whether y_i, the number of successful coin tosses reported by a subject, is within the respective sets. The main independent variable MACHINE is a dummy which indicates whether a subject reported to a machine (FORM). The two treatments with human interaction (CALL and CHAT) serve as the reference category. "Base rate" refers to the proportion of positive outcomes for the dependent variable which the regression model predicts for the reference category. "Expected rate" refers to the outcome for the dependent variable that is expected under truthful reporting. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Only data from Wave 1 are used. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B.3. Suspicious vs. credible outcomes across MACHINE and HUMAN treatments in Wave 2

	(1)	(2)	(3)
	Panel (a): Suspicious over-reporting		
Dependent variable:	$y_i \in \{8, 9, 10\}$	$y_i \in \{7, 8, 9, 10\}$	$y_i \in \{9, 10\}$
MACHINE	0.165***	0.213***	0.091***
	(0.042)	(0.061)	(0.033)
Base rate	0.046*	0.192***	0.032
	(0.025)	(0.047)	(0.020)
Expected rate	0.055	0.172	0.011
	Panel (b): Credible over-reporting		
Dependent variable:	$y_i \in \{6, 7\}$	$y_i \in \{6\}$	$y_i \in \{6, 7, 8\}$
MACHINE	0.029	-0.036	0.097
	(0.073)	(0.062)	(0.073)
Base rate	0.384***	0.242***	0.399***
	(0.060)	(0.0522)	(0.060)
Expected rate	0.322	0.205	0.367
Controls:			
Subject characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	2	2	2
Observations	211	211	211

Notes: Probit average marginal effects with robust standard errors in parentheses. The dependent variable is a dummy which indicates whether y_i, the number of successful coin tosses reported by a subject, is within the respective sets. The main independent variable MACHINE is a dummy which indicates whether a subject reported to a machine (FORM and ROBOT). The two treatments with human interaction (CALL and CHAT) serve as the reference category. "Base rate" refers to the proportion of positive outcomes for the dependent variable which the regression model predicts for the reference category. "Expected rate" refers to the outcome for the dependent variable that is expected under truthful reporting. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Only data from Wave 2 are used. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B.4. Risk aversion and cheating by treatments

	(1)	(2)	(3)
Dependent variable	$y_{it} = 1$: coin toss reported as successful		
FORM	0.080***	0.081***	0.080***
	(0.019)	(0.019)	(0.019)
ROBOT	0.069***	0.069***	0.071***
	(0.025)	(0.025)	(0.026)
CHAT	0.017	0.017	0.016
	(0.022)	(0.022)	(0.021)
Risk aversion		-0.003	-0.016
		(0.009)	(0.014)
Risk aversion × FORM			0.031
			(0.021)
Risk aversion × ROBOT			0.017
			(0.027)
Risk aversion × CHAT			0.003
			(0.023)
Controls			
Subject Characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Wave	1&2	1&2	1&2
Observations	4,680	4,680	4,680
Subjects	468	468	468

Notes: Probit average marginal effects with robust standard errors, corrected for clustering at the individual level, in parentheses. The dependent variables are a dummy indicating whether subjects 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 the reference category. The risk aversion measure is based on subjects' 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 recoded this measure such that larger values indicate higher risk aversion and then normalized it so that the variable "Risk aversion" has a mean of zero and a standard deviation of one. Control variables include subjects' age in years and dummies for gender, Swiss citizenship, fields of study, and experimenters. Data from Wave 1 and Wave 2 are pooled. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B.5. Order effects on reporting (Experiment 1)

	(1)
Dependent variable	$y_{it} = 1$: coin toss reported as successful
MACHINE	0.072***
	(0.016)
Period 2	-0.000
	(0.033)
Period 3	0.025
	(0.033)
Period 4	0.030
	(0.031)
Period 5	-0.007
	(0.030)
Period 6	0.021
	(0.032)
Period 7	0.006
	(0.032)
Period 8	-0.002
	(0.031)
Period 9	0.017
	(0.031)
Period 10	0.028
	(0.030)
Controls:	
Field of study	yes
Experimenter FE	yes
Observations	4,680
Subjects	468

Notes: Probit average marginal effect with standard errors, corrected for clustering at the individual level, in parentheses. The dependent variable is a dummy indicating whether a subject 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 (FORM and ROBOT). The two treatments with human interaction (CALL and CHAT) serve as the reference category. "Period t" dummies indicate that a report was the t-th outcome (out of 10) which a subject reported. Control variables include a subject's age and dummies for gender, Swiss citizenship, fields of study, and for the experimenters' identities. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table B.6. Cheating and social distance cues (with interactions)

	(1)	(2)	(3)
Dependent variable	$y_{it} = 1$: coin toss reported as successful		
CALL	-0.075*	-0.128***	-0.073**
	(0.042)	(0.047)	(0.034)
Same gender	-0.007		
	(0.047)		
Same gender × CALL	0.025		
	(0.058)		
Same native language		-0.060	
		(0.047)	
Same native language × CALL		0.094	
		(0.057)	
Same gender & native language			-0.001
			(0.047)
Same gender & native language × CALL			0.033
			(0.059)
Controls:			
Subject characteristics (w/o Swiss)	yes	yes	yes
Experimenter FE	yes	yes	yes
Data from	$CALL_2$	$CALL_2$	$CALL_2$
	$FORM_2$	$FORM_2$	$FORM_2$
Observations	1,420	1,420	1,420
Subjects	142	142	142

Notes: Probit average marginal effects with robust standard errors, corrected for clustering at the individual level, in parentheses. The dependent variable is a dummy indicating whether a subject reported a coin toss as successful (10 observations per subject). The main independent variables are (i) a dummy indicating whether observations are from treatment CALL (as opposed to FORM, the baseline) in Wave 2, (ii) dummies indicating whether the subject and the experimenter had the same gender, or native language (Swiss German), or both features the same as the experimenter, and (iii) the interaction of these dummies. Control variables include subjects' age in years and dummies for gender, fields of study, and experimenters. The dummy for Swiss citizenship was omitted to avoid co-linear regressors: All experimenters were Swiss-German speakers and only 4 subjects in the relevant treatments were Swiss-German speakers but not Swiss citizens. The data used are always from treatment CALL and FORM in Wave 2. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table B.7. Cheating and social distance cues (without interactions)

	(1)	(2)	(3)
Dependent variable	$y_{it} = 1$: coin toss reported as successful		
FORM	0.080***	0.083***	0.082***
	(0.023)	(0.024)	(0.024)
ROBOT	0.071***	0.069***	0.068***
	(0.026)	(0.026)	(0.026)
CHAT	0.017	0.018	0.017
	(0.022)	(0.022)	(0.022)
Same gender	-0.011		
	(0.017)		
Same native language		0.026	
		(0.027)	
Same gender & native language			0.019
			(0.019)
Controls:			
Subject characteristics	yes	yes	yes
Experimenter FE	yes	yes	yes
Data from	Wave 1&2 w/o FORM_2	Wave 1&2 w/o FORM_2	Wave 1&2 w/o FORM_2
Observations	3,930	3,930	3,930
Subjects	393	393	393

Notes: Probit average marginal effects with robust standard errors, corrected for clustering at the individual level, in parentheses. The dependent variable is a dummy indicating whether a subject reported a coin toss as successful (10 observations per subject). The main independent variables are (i) dummies which indicate whether a subject was in either treatments FORM, ROBOT, CHAT (CALL is the reference category), (ii) dummies indicating whether the subject and the experimenter had the same gender, or native language (Swiss German), or both features the same as the experimenter. Control variables include subjects' age in years and dummies for gender, fields of study, Swiss citizenship, and experimenters. The data used are always from Wave 1 and Wave 2, but without data from FORM in Wave 2. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table B.8. Reporting behavior in Part B of Experiment 2

	(1)
Dependent variable:	$y_i = 1$ coin toss reported as successful
Choice = Form	0.084**
	(0.033)
Choice = Call (Baseline)	0.584***
	(0.025)
Controls: Subject characteristics	yes
Observations	880
Subjects	88

Probit average marginal effects with robust standard errors in parentheses. The dependent variable is a dummy indicating whether a subject reported a coin toss as successful (10 observations per subject). The main independent variable is a dummy which indicates whether a subject chose to report via form (choice for reporting via call is the level predicted by the model in the reference category). Control variables include subjects' age in years and dummies for gender, Swiss citizenship, and for different fields of study. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

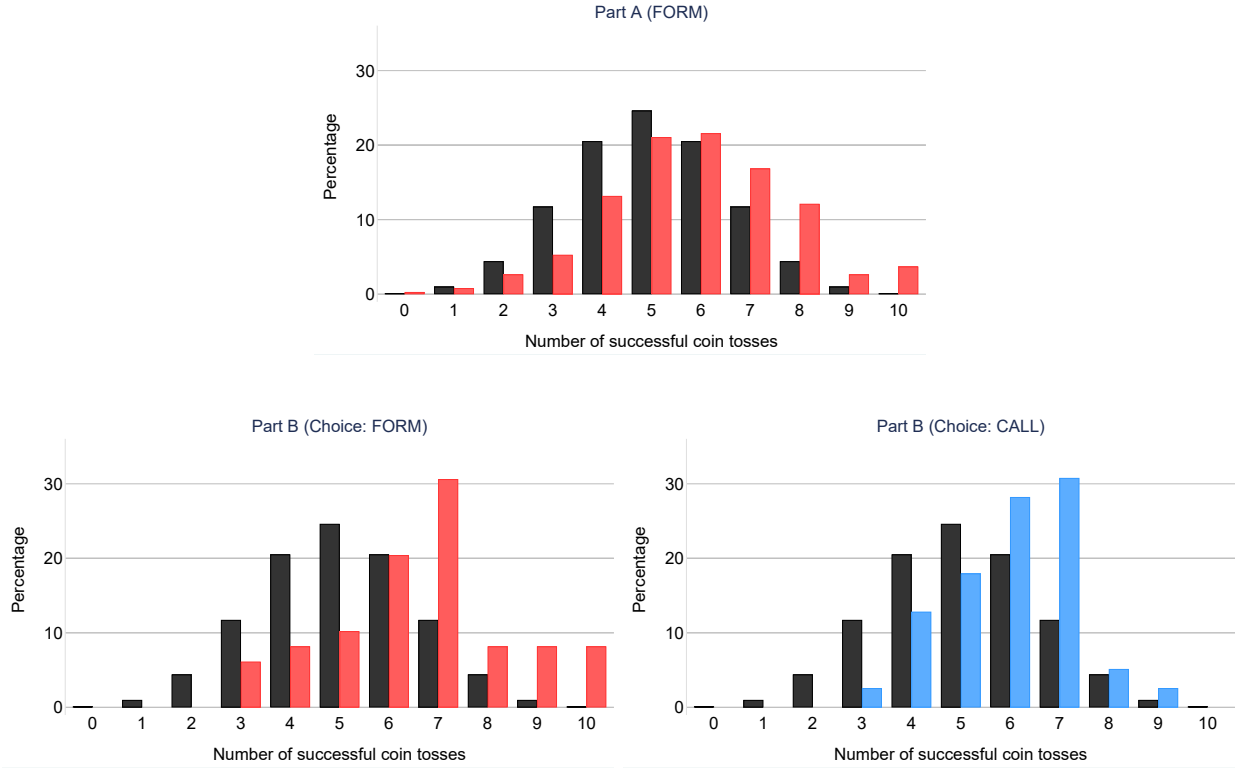
In the first wave of the main experiment we included a measure of social distance (the "Inclusion of Other in the Self" scale, see Aron et al., 1992). Subjects had to indicate how close they felt to the experimenter by selecting one out of five pairs of circles that varied by how much they overlap (which serves as an indicator of closeness). Table B.9 shows that there are no significant differences in perceived social distance across conditions (the smallest p-value is 0.532). However, we do not think that this is a good proxy of perceived social distance for our purposes because asked subjects about their feelings of closeness with the experimenter in general, rather than during the reporting stage. Yet, the time aspect is crucial here because at the beginning of the experiment subjects interacted with the experimenter over Skype in all treatments. Thus, if subjects evaluated their entire experience in the experiment rather than only the reporting stage when answering this question, it is not surprising that we do not find any difference in perceived social distance across treatments.

Table B.9. Social distance effects

	(1)
Dependent variable:	Self-reported social distance (0 to 4)
FORM	-0.006
	(0.137)
CHAT	-0.084
	(0.135)
Constant	3.620***
	(0.323)
Controls:	
Subject characteristics	yes
Experimenter FE	yes
Wave	1
Observations	257

Notes: OLS estimates with robust standard errors in parentheses. The dependent variable is the self-reported distance to experimenter (0 to 4, based on the "Inclusion of Other in the Self" scale). The main independent variables are dummies which indicate whether a subject was in treatment FORM or CHAT of Wave 1 (CALL is the reference category). Control variables include subjects' age in years and dummies for gender, fields of study, Swiss citizenship, and experimenters. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Figure B.1. Distribution of successful coin tosses in Part A and Part B of Experiment 2

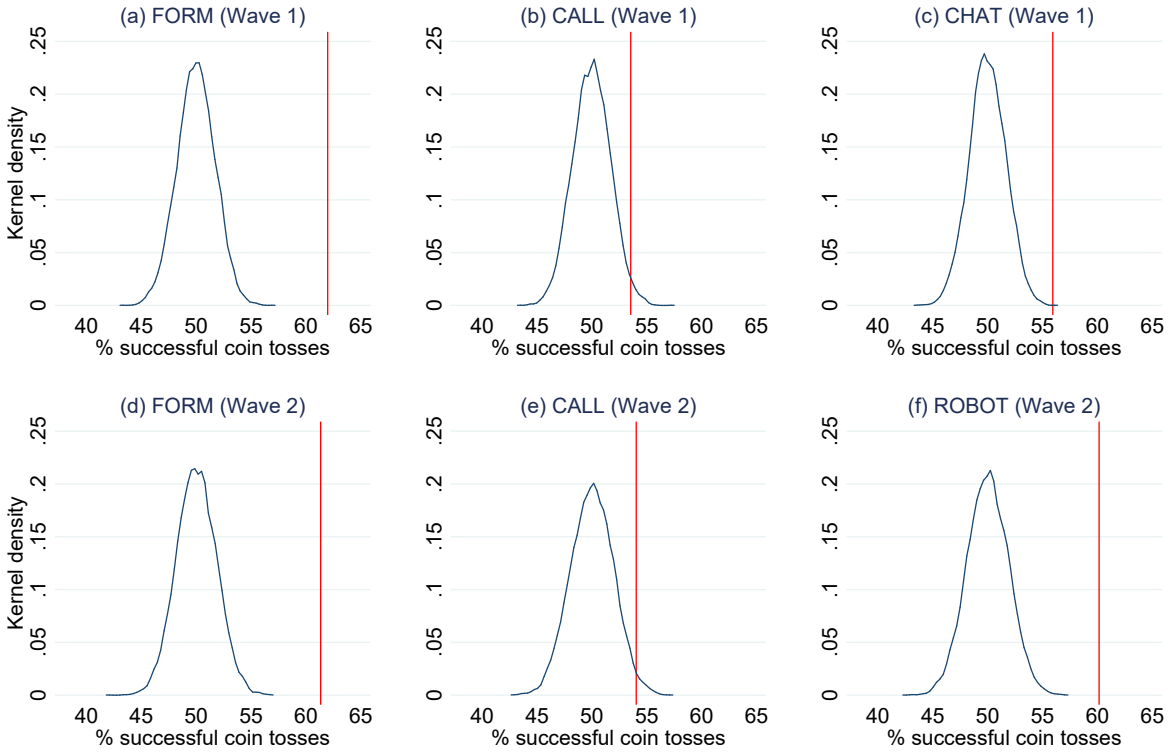


Notes: Colored bars depict actual observations by reporting channel; blue=choice to report via call, red=choice to report via form, and black bars depict the distribution expected under truthful reporting.

Appendix C: Simulation Analysis

Empirical distributions of successful coin tosses may deviate from their theoretical counterpart (i.e., binomial distribution), even if the coins are fair and everyone reports their outcomes truthfully. Due to random fluctuations, actual frequencies of successful coin flips may not exactly match the expected frequencies. In this section, we explore with simulations whether the observed treatment effects can, in principle, be explained by random fluctuations. To this end, we simulated 10,000 coin flipping experiments for each treatment with the same number of subjects and coin flips as in the respective treatments. In the simulations, we assume that each coin toss is generated by a binomial process with an underlying success rate of 50% (i.e., truthful reporting).

Figure C.1. Simulated and observed success rates by treatment and wave (Experiment 1)



Notes: Panels (a) to (f) show kernel densities of the percentages of successful coin flips resulting from 10,000 simulated samples assuming n×10 truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the different treatment groups (n=86 for FORM in Wave 1; n=75 for FORM in Wave 2; n=85 for CALL in Wave 1; n=67 for CALL in Wave 2; n=86 for CHAT; n=69 for ROBOT). The red vertical lines represent the average percentages of successful coin flips in the actual data.

Our first observation is that it is unlikely that subjects reported their outcomes completely honestly, no matter which treatment or wave. Panels (a) to (f) in Figure C.1 show that the actual success rates (vertical red lines) all lie outside (for FORM and ROBOT) or at the right tail (for CALL and CHAT) of the distributions of simulated success rates. If all subjects reported truthfully, then the probability of observing the same or a larger success rate as in our experiments are p=0.0000 (for FORM in waves 1 and 2, as well as ROBOT), p=0.0210 (for CALL in Wave 1), p=0.0197 (for CALL in Wave 2), and p=0.0001 (for CHAT).[2] This suggests that subjects, on average, cheated in each condition and wave, at least to some degree.

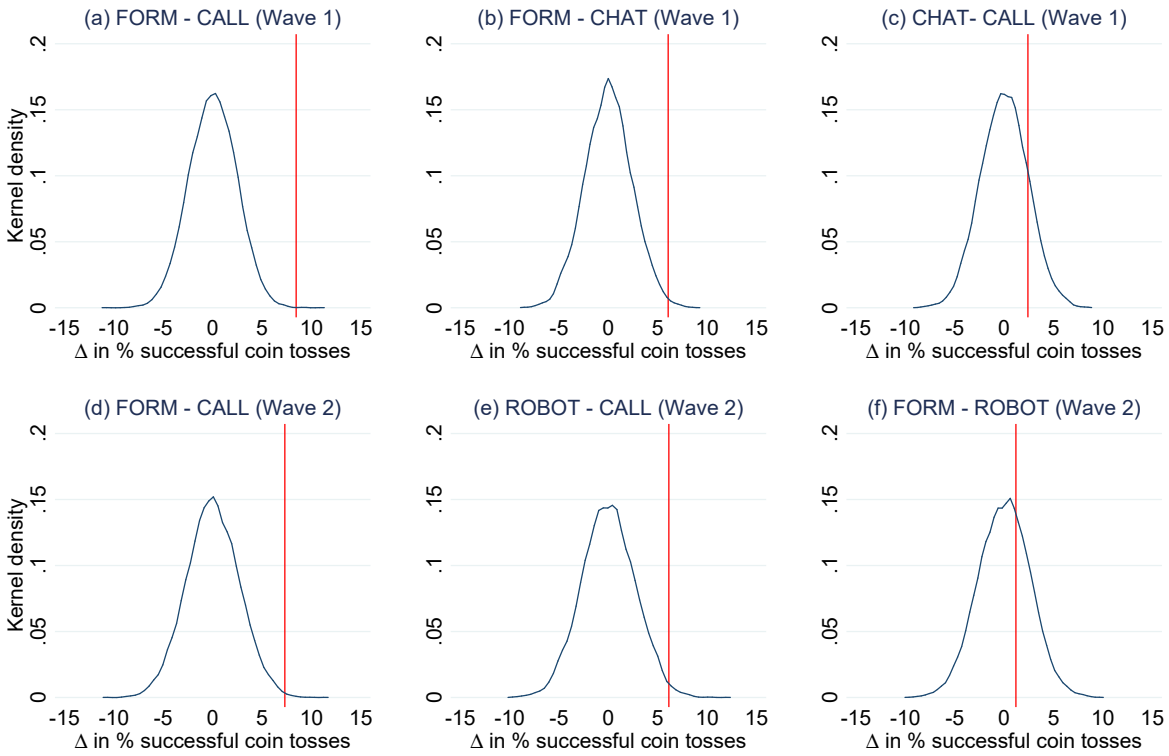
We next focus on treatment differences, and ask whether they could have occurred simply by chance. Figure C.2 presents the distributions of simulated treatment effects (assuming that all subjects reported truthfully) and the actual treatment effects represented by the red vertical lines. The probability that participants behaved honestly but generated the observed absolute differences in success rates between FORM and CALL are p=0.0006 and p=0.0050 for waves 1 and 2, respectively (panels a and d).[3] Similarly, the probabilities of observing the same or larger absolute treatment differences are p=0.0117 for FORM versus CHAT and p=0.0235 for ROBOT versus CALL, respectively (panels b and e). Given our sample size, it is thus unlikely that the treatment differences between human and machine conditions are due to random fluctuations in coin tossing. In contrast, the corresponding probabilities are p=0.3214 for the difference between human conditions (CHAT and CALL), and p=0.6534 for the difference between machine conditions (FORM and ROBOT), respectively (panels c and f).

We further use the simulated data to complement our analysis of suspicious versus credible reporting (see "Mechanism" section in the paper). Panels (a) and (b) in Figure C.3 show that it is unlikely that the share of people reporting suspicious outcomes (i.e., 8 or more successful coin flips) in MACHINE and HUMAN are driven by honest reporting (p=0.0000 and p=0.0365, respectively). Similarly, the simulations in panels (c) and (d) show that the reported percentages of credible outcomes (i.e., 6 or 7 successful coin flips) is hard to reconcile with honest reporting (p=0.0017 for MACHINE and p=0.0031 for HUMAN). Panel (e) shows that the likelihood of observing the same or a larger treatment difference in suspicious reporting between MACHINE and HUMAN is zero (p=0.0000). In contrast, panel (f) shows that the actual and simulated treatment differences for credible outcomes largely overlap (p=0.8313).

[2]We report one-sided, simulated p-values for all tests of honest behavior within treatments because we assume that people do not cheat to their disadvantage.

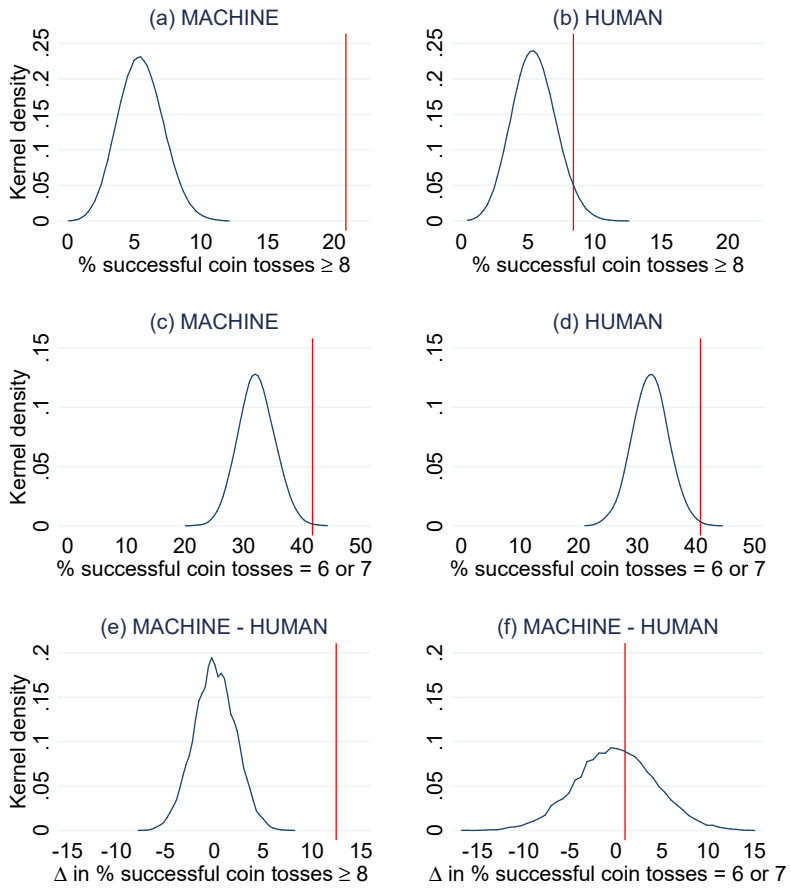
[3]We report two-sided simulated p-values for all tests that compare differences across treatments.

Figure C.2. Simulated and observed treatment differences (Experiment 1)



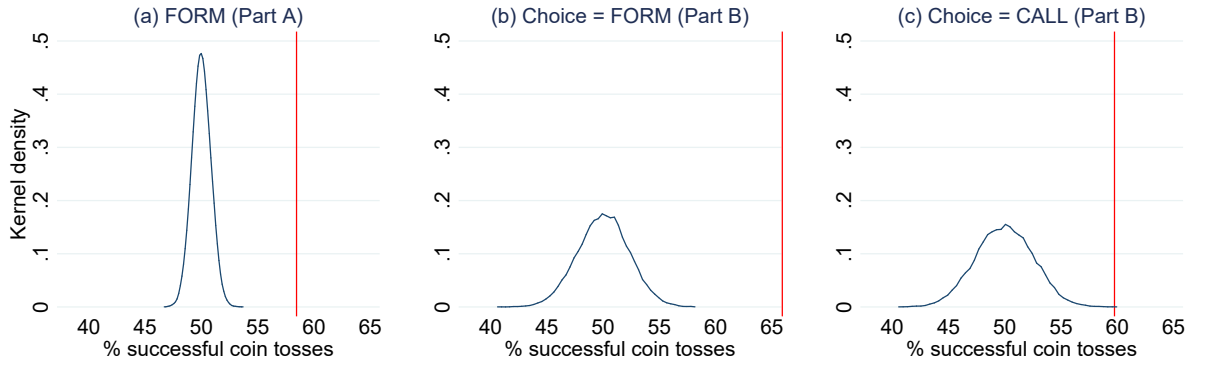
Notes: Panels (a) to (f) display kernel densities for the differences in successful coin flips between treatments for each of the 10,000 simulated samples, assuming n$\times$10 truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the different treatment groups (n=86 for FORM in Wave 1; n=75 for FORM in Wave 2; n=85 for CALL in Wave 1; n=67 for CALL in Wave 2; n=86 for CHAT; n=69 for ROBOT). The red vertical lines indicate treatment differences in the actual data.

Figure C.3. Simulation analysis of suspicious and credible outcomes (Experiment 1)



Notes: Panels (a) and (b) (respectively, c and d) show kernel densities of the percentages of 8 or more (respectively, 6 or 7) successful coin flips resulting from 10,000 simulated samples assuming n$\times$10 truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the different conditions (n=230 for MACHINE; n=238 for HUMAN). The red vertical lines represent the share of 8 or more (respectively, 6 or 7) successful coin tosses that are observed in the actual data. Panels (e) and (f) display kernel densities for the simulated differences in 8 or more (respectively, 6 or 7) successful coin flips between MACHINE and HUMAN, assuming truthfully reported coin tosses. The red vertical line indicate observed freqeunces and treatment differences in the actual data.

Figure C.4. Simulation analysis or Part A and B (Experiment 2)



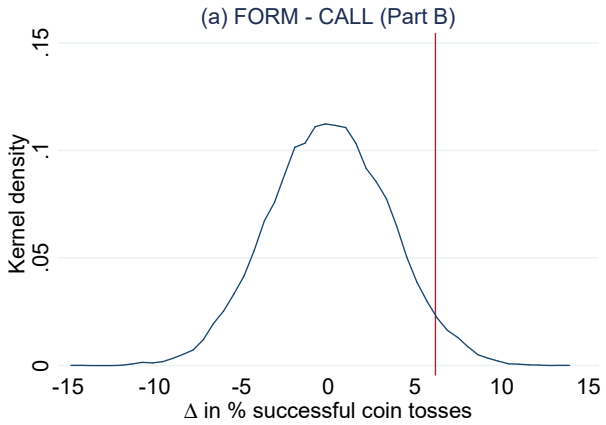
Notes: Panels (a) to (f) display kernel densities for the differences in successful coin flips between treatments for each of the 10,000 simulated samples, assuming n$\times$10 truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the different treatment groups (n=86 for FORM in Wave 1; n=75 for FORM in Wave 2; n=85 for CALL in Wave 1; n=67 for CALL in Wave 2; n=86 for CHAT; n=69 for ROBOT). The red vertical lines indicate treatment differences in the actual data.

Finally, we ran a similar simulation analysis for Experiment 2, simulating another 10,000 coin flipping experiments with honest reporting for Part A and B with the same number of subjects and coin flips as in Parts A and B. The simulated distributions in panels (a) to (c) of Figure C.4 again highlight that actual success rates lie outside or at the right tail of the simulated distributions (p=0.0000 for FORM in Parts A and B, p=0.0001 for CALL in Part B). This suggests that subjects also cheated in Parts A and B of Experiment 2.

The likelihood that participants behaved honestly but generated the observed absolute difference in success rates between those who chose FORM and those who chose CALL is 0.0726 (see Figure C.5). Note that the sample in Part B is less than a quarter of the observations in Part A as subjects were invited with probability of 25% (and 12% of the invited subjects did not show up for Part B). We did so because Experiment 2 was designed to study the selection decision by participants in Part A rather than their cheating behavior in Part B.

In sum, the simulation analysis demonstrates that the law of large numbers applies to our specific sample sizes and that our main results are not just the result of random fluctuations in coin tossing.

Figure C.5. Simulated and observed differences between chosen reporting channels (Experiment 2)



Notes: This figure displays the kernel density for the differences in successful coin flips between samples for each of the 10,000 simulated samples, assuming n$\times$10 truthfully reported coin tosses (i.e., a probability rate of 50%), where n corresponds to the actual sample size of the different groups (n=39 for subjects who chose CALL in Part B; n=49 for subjects who chose FORM in Part B). The red vertical lines indicate treatment differences in the actual data.

Appendix D: Survey experiment

Design: We conducted a survey experiment on Amazon Mechanical Turk (MTurk) to provide additional evidence on the underlying mechanism. In particular, we explore how subjects perceive the reporting stage in treatments CALL and ROBOT in terms of human presence and social image concerns.

We explained to the Mturk subjects that we had conducted an experiment and that their task was to take the perspective of the participants in the experiment. Subject then read a detailed description of Experiment 1 (Wave 2). Specifically, we informed them that participants in Experiment 1 were individually welcomed by "a person who carried out the experiment" (referred to as the "other person") via Skype chat. They also learned that after the welcome stage, participants received a link to an online survey in which they were instructed to perform the coin tossing task.[4] Subsequently, we explained that there were two conditions – treatments CALL and ROBOT – and how they differed with regards to reporting the outcomes of the coin tosses.

Following the description of the original experiment, Mturk subjects had to report how they would feel as a participant in the experiment. First, they answered four questions capturing their perceptions of human presence (i.e., feelings of closeness in terms of socially interacting with the other person) for one of the two treatments. Specifically, we asked them the following questions (on a Likert scale ranging from 1 "Not at all" to 7 "Very much"):

As you report the outcomes of your coin tosses to the [other person on the Skype call... / voice response system (that uses the pre-recorded voice of the other person)...

1. *How close would you feel to the other person?*
2. *How strongly would you feel the presence of the other person?*
3. *How connected would you feel to the other person?*
4. *To what extent would you feel that you are alone?*

We then elicited social image concerns for the same treatment using the following three questions (on a Likert scale ranging from 1 "Not at all" to 7 "Very much"):

[4] For simplicity, we slightly modified the description of the coin tossing task such that reporting HEADS would always yield US$2 (about CHF 2) and TAILS nothing for each of the ten tosses.

5. *How concerned would you be about what the other person thinks about you?*

6. *How much would you care about leaving a good impression on the other person?*

7. *How important would it be for you that the other person thinks you are honest?*

After subjects answered those seven questions for one of the two treatments (e.g., CALL) they were shown the same set of questions for the other treatment (e.g., ROBOT). We randomized the order of the treatments across subjects.[5] We also randomized the order in which the questions appeared within each block of questions and held the specific sequence within a block constant across treatments.

Next, we asked subjects to guess the average number of successful coin tosses reported in each treatment. We incentivized their predictions by paying a bonus of \$1 for the 20% most accurate subjects.

Finally, subjects provided information on their socioeconomic background (age, gender, education, employment status, and relative income). They could earn another bonus of \$0.5 if they passed a final attention check (i.e., they had to compute a participant's earnings resulting from a randomly given number of successful coin tosses).

Procedures: The survey experiment took place in June 2020. Subjects were recruited via Amazon Mechanical Turk (MTurk). They received a base payment of \$0.5 for their participation. They were required to have an approval rate of 98% or more and at least 100 HITs (tasks on MTurk) completed. In order to participate, they had to pass an initial attention check. In total, we collected responses from n=156 subjects who gave consent to participate and passed all attention checks. On average, it took about 10 minutes for subjects to complete the survey.

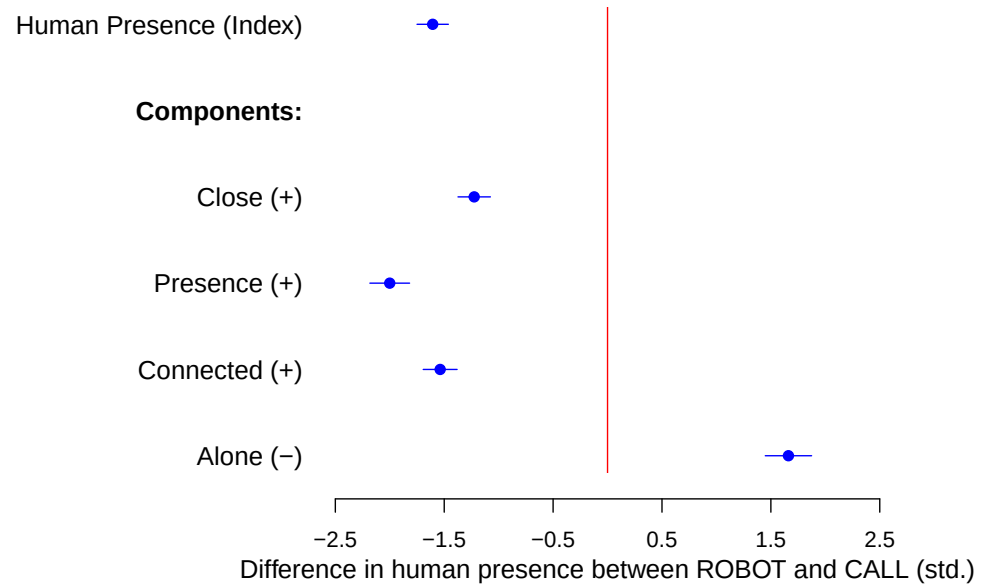
Results: We standardized the responses for each question 1 to 4 using the mean and standard deviation in the ROBOT condition. We then created a human presence index using the unweighted average of the standardized responses.[6] We followed the same procedure to construct an index of social image concerns using responses to questions 5 to 7.

Figure D.1 shows the difference between ROBOT and CALL in perceived human presence, both with respect to the index and the individual questions. The results reveal large differences in perceived human

[5] The results are largely the same if we only use the data from the first treatment and compare responses between subjects.

[6] We reverse-coded responses to question 4 as "feeling alone" indicates less human presence.

Figure D.1. Human Presence in ROBOT relative to CALL



Notes: Differences in standardized responses between treatments ROBOT and CALL. The signs in parentheses denote whether the components were positively or negatively coded for the construction of the human presence index. Original responses are based on Likert scales ranging from 1 ("Not at all) to 7 ("Very much") and were standardized using the mean and standard deviation of the corresponding question in ROBOT. The index is the unweighted average of the following four components (with the component "alone" being reverse coded): "*How close would you feel to the other person?*" (Close), "*How strongly would you feel the presence of the other person?*" (Presence), "*How connected would you feel to the other person?*" (Connected), "*To what extent would you feel that you are alone?*" (Alone). Error bars denote the standard error of the mean.

presence across conditions. For example, the human presence index is about 1.6 standard deviations lower in ROBOT than in CALL. The results are similar in magnitude for the individual components of the index.

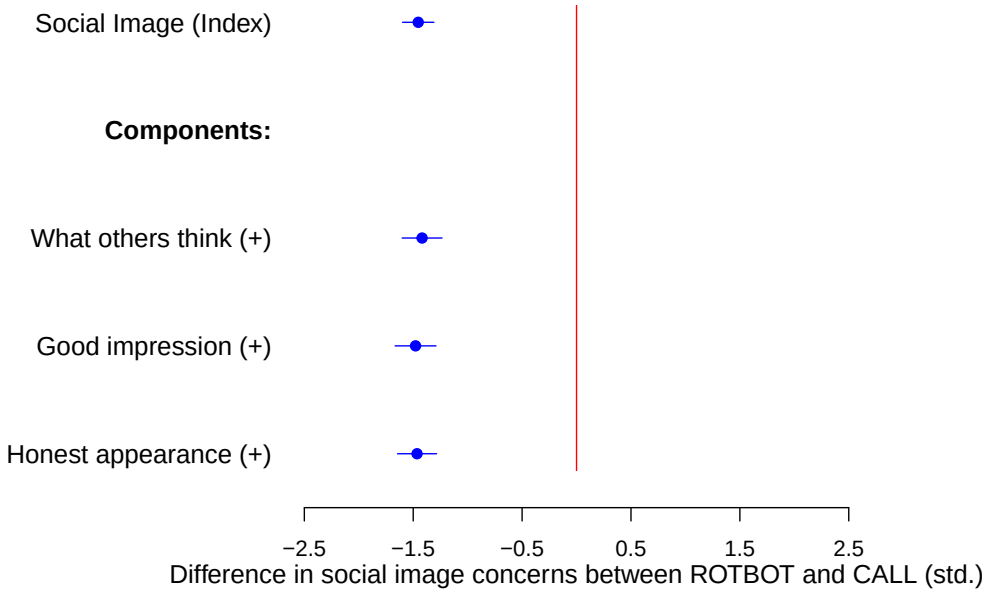
Figure D.2 shows that subjects were substantially less concerned about their social image in ROBOT than in CALL. The score of the social image index is roughly 1.5 standard deviations lower in ROBOT. The results are again remarkably consistent across each individual component of the index.

To investigate the extent to which human presence moderates differences in social image concerns between ROBOT and CALL, we estimate the following OLS regression model:

$$SocialImage_{it} = \alpha + \beta_1\, ROBOT_t + \beta_2\, HumanPresence_{it} + \gamma\, X_i + \epsilon_{it},$$

where the dependent variable $SocialImage_{ic}$ is the score of the social image index of subject i in treatment t. $ROBOT_t$ is a dummy that takes on a value of one for responses in treatment ROBOT,

Figure D.2. Social Image Concerns in ROBOT relative to CALL



Notes: Differences in standardized responses between treatments ROBOT and CALL. The signs in parentheses denote whether the components were positively or negatively coded for the construction of the social image index. Original responses are based on Likert scales ranging from 1 ("Not at all) to 7 ("Very much") and were standardized using the mean and standard deviation of the corresponding question in ROBOT. The index is the unweighted average of the following three components: "*How concerned would you be about what the other person thinks about you?*" (What others think), "*How much would you care about leaving a good impression on the other person?*" (Good impression), "*How important would it be for you that the other person thinks you are honest?*" (Honest appearance). Bars denote standard error of the mean.

and zero for CALL. $HumanPresence_{it}$ is the score of the human presence index of subject i in treatment t, and X_i is a vector of variables controlling for subjects' socioeconomic background. Because we have two observations for each subject we cluster standard errors at the subject level.

Column 1 of Table D.1 shows the unconditional effect of ROBOT on social image concerns. The social image index is 1.5 standard deviations lower in ROBOT relative to CALL ($p<0.001$, t-test). In column 2, we add the human presence index as an additional explanatory variable. The results show a strong positive correlation between human presence and social image concerns—a one standard deviation increase in the human presence index is associated with a 0.7 standard deviations increase in social image concerns ($p<0.001$, t-test). Adding the human presence index increases the model's explanatory power, as measured by its R^2, and it reduces the magnitude of the coefficient of ROBOT by 81.4%. As a result, the coefficient of ROBOT is no longer significant ($p=0.116$, t-test). A Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973) yields very similar results. A large portion (88.6%) of the

Table D.1. Decomposing image concerns differences between ROBOT and CALL

	(1)	(2)	(3)	(4)
Dependent variable		Social Image (Index)		
ROBOT	-1.454***	-0.270	-1.454***	-0.269
	(0.114)	(0.171)	(0.109)	(0.173)
Human Presence (Index)		0.737***		0.738***
		(0.071)		(0.072)
Constant	0.000	0.000	-0.193	-0.175
	(0.074)	(0.070)	(0.475)	(0.396)
% explained (Blinder-Oaxaca)		88.6		88.2
Controls	no	no	yes	yes
R^2	0.344	0.563	0.400	0.594
Observations	312	312	312	312

Notes: OLS estimates with standard errors clustered at the individual level in parentheses. The dependent variable is the standardized index for perceived social image concerns. "ROBOT" is a dummy variable that takes a value of one for treatment ROBOT, and zero for CALL. "Human Image (Index)" is the standardized human presence index. Control variables include subjects' age in years, dummies for gender, education, employment status and relative income. The results of the Blinder-Oaxaca decomposition analysis is shown at the bottom of the table. The numbers reveal how much how much of the treatment effect on social image concerns is explained by differences in the human presence index. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

impact of ROBOT on social image concerns can be explained by differences in perceived human presence between the two treatments. The results remain unchanged when we control for subjects' background characteristics (columns 3 and 4). Together, the results from the additional survey experiment are consistent with a mechanism based on social image concerns. Subjects experience a lower sense of human presence when reporting to a machine rather than a person, which in turn lowers their desire to be perceived as an honest person.

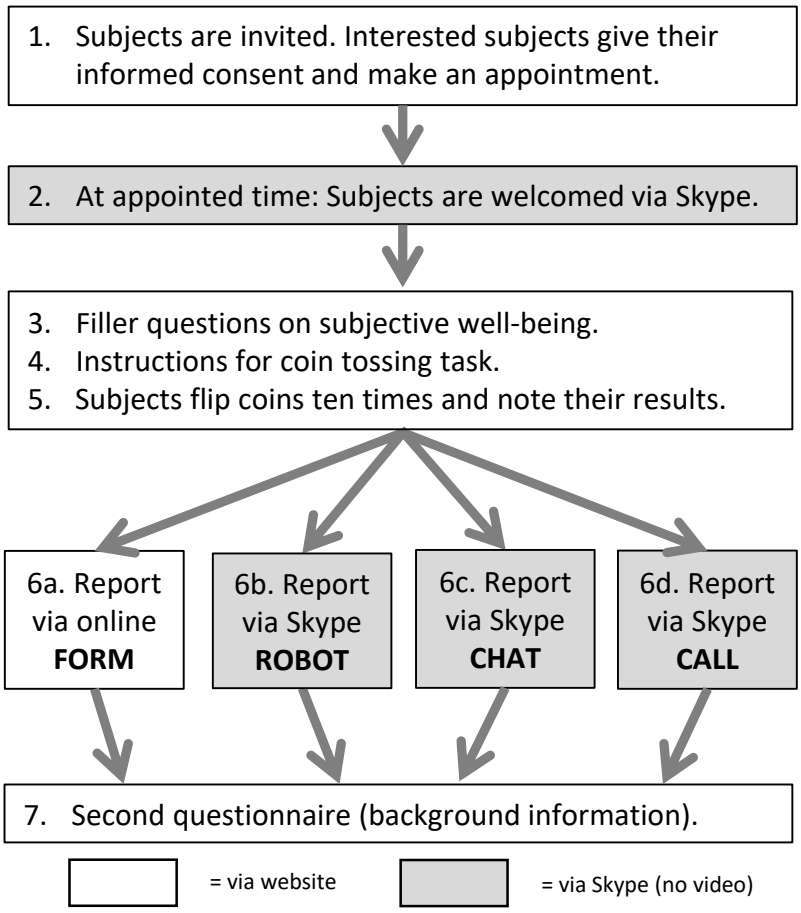
Appendix E: Procedures and instructions for Experiment 1

The experiment had the following structure: In step 1, subjects from the university's subject pool were invited (via email) to participate.[7] We excluded psychology students as they often participate in experiments that involve deception as well as individuals who previously participated in an experiment on lying or cheating. Moreover, we only recruited subjects who had participated at least once in an economic lab experiment to ensure that they trusted our instructions and payment procedure. Invited subjects were told that the study would require a Skype account and that they would need to provide their address as their earnings would be mailed to them. If subjects gave consent, they could register on our website and make an appointment. At the scheduled time, they were contacted by an experimenter via Skype in step 2.

The experimenter welcomed the subjects and sent them, via Skype's chat function, a personalized link to an online questionnaire. The first part of the questionnaire contained filler questions on subjective well-being and life satisfaction (step 3). In step 4, subjects received the instructions for the coin tossing task and were informed how to report the outcomes. Subsequently, subjects flipped a coin ten times and noted the results (step 5). When ready, subjects could proceed to report the outcomes in step 6. In it, they then reported via an online form (treatment FORM), Skype call (treatment CALL), Skype chat (treatment CHAT), or automated voice response system (treatment ROBOT). Finally, they received a link for an exit questionnaire which elicited additional information, including their address to which their earnings were sent. Figure E.1 visualizes these steps:

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

Figure E.1. Timeline of Experiment 1



Notes: Wave 1 featured steps 6a, 6c, and 6d; Wave 2 featured steps 6a, 6b, and 6c.

The following pages display the instructions (translations from German) for the coin tossing task. Content in frames was shown in the subjects' web browser.

Step 4: Instructions for the coin tossing task

Now you can win money!

All participants of this study can earn up to **CHF 20.-** *The amount you earn depends on what you report. It is thus very important that you read the instructions carefully.*

Please have the **coin, the paper, and the pencil** *ready. You will now be asked to toss the coin ten times and to note the results (heads or tails) on paper. Using the corresponding payment tables, you can see whether you won the toss or not. Each win increases your income by CHF 2.-, meaning that you can earn up to CHF 20.-*

[in treatment FORM]

You will be asked later to report the results of your coin tosses in writing on one of the following pages.

[in treatment CALL]

You will be asked later to report the results of your coin tosses orally to the person with whom you spoke early via Skype call (without video).

[in treatment CHAT]

You will be asked later to report the results of your coin tosses in writing to the person with whom you spoke early via Skype chat.

[in treatment ROBOT]

You will be asked later to report the results of your coin tosses orally via Skype on an answering machine (without video).

Example of a payment table :

		
Yield	CHF 2.-	CHF 0.-

You win if you toss tails, and your income is increased by CHF 2.–

If you toss heads, you lose and will not earn anything more.

Please click on "continue" to continue.

[continue-button]

Step 5: Coin tosses

Please pick up the coin and toss it 10 times. **Please note the corresponding result for each toss, i.e., heads or tails.** *You can see whether you won on the corresponding payment table below.*

[in treatment FORM]

You will be asked later to report the results of your coin tosses in writing on one of the following pages.

[in treatment CALL]

You will be asked later to report the results of your coin tosses orally to the person with whom you spoke earlier via Skype call.

[in treatment CHAT]

You will be asked later to report the results of your coin tosses in writing to the person with whom you spoke earlier via Skype chat.

[in treatment ROBOT]

You will be asked later to report the results of your coin tosses orally via Skype on an answering machine.

Please begin now with toss 1 and write down the results of the remaining 9 tosses.

Toss 1

		
Yield	CHF 0.-	CHF 2.-

...

[The corresponding pictures for tosses 2 to 9 are omitted. The sequence was "Heads" /"Tails" /"Tails" /"Heads" /"Tails" /"Tails" /"Heads" /"Heads" .]

Toss 10

		
Yield	CHF 2.-	CHF 0.-

Please click on "continue" when you have noted all 10 tosses.

[continue-button]

Step 6 (i): Preparation to report

In treatment FORM:

> *You will now be asked to report the results of your coin tosses. Please only answer with "heads" or "tails". You do not need to provide any explanation for your responses.*
> *Please click on "continue" to continue.*
>
> [continue-button]

In treatment CALL:

> *Please call back by Skype. Thank you.*
> *You can close this window when the connection has been made.*

Once subjects called back the experimenter welcomed them with the following spoken text:
Welcome back! We now ask you to report the results of your coin tosses. Please only answer with "heads" or "tails". You do not need to provide any explanation for your responses.

In treatment CHAT:

> *Please contact us by Skype chat. Thank you.*
> *You can close this window when the connection has been made.*

Once subjects wrote back the experimenter welcomed them by writing the same text as the one spoken in CALL (see above).

In treatment ROBOT:

> *We now ask you to report the results of your coin tosses orally via Skype on an answering machine.*
> *Please only answer with "heads" or "tails". You do not need to provide any explanation for your responses.*
> *Please have the results of your coin tosses ready. Afterwards you can leave us your address on an internet page so that we can send you your earnings*
> [The instructions on how to call the specific Skype contact with the answering machine is omitted.]
> *You can close this window when the connection has been made.*

Once subjects were connected to the answering machine, it welcomed them by using one of the experimenter's pre-recorded greeting with the same text as in CALL (see above).

Step 6 (ii): Reporting

In treatment FORM:
There were ten separate screens for each of the ten coin tosses. Each screen elicited the response via a text entry field which only accepted the German equivalents for "Heads" or "Tails" (case-insensitive). Below, we show the screen which elicits the response for coin toss 1, the other nine screens are analogous:

You will win CHF 2 in toss 1 if you have "tails". Did you have "heads" or "tails"?
[Text entry box]
Please click on "continue" once you have entered your result. [continue-button]

After having reported their result for all ten coin tosses, subjects were forwarded to another screen with the exit-questionnaire.

In treatment CALL:
The experimenter orally asked subjects exactly the same question as in the above example screen for treatment FORM to report the outcome of each coin toss, and subjects answered orally. After having reported their result for all ten coin tosses, the experimenter sent subjects a link for the exit-questionnaire via Skype chat.

In treatment CHAT:
The experimenter asked subjects in writing exactly the same question as in the above example screen for treatment FORM to report the outcome of each coin toss, and subjects answered in writing. After having reported their result for all ten coin tosses, the experimenter sent subjects a link for the exit-questionnaire via Skype chat.

In treatment ROBOT:
The pre-recorded experimenter's voice asked subjects exactly the same question as in the above example screen for treatment FORM to report the outcome of each coin toss, and subjects answered orally.
As in the other treatments, subjects in ROBOT received a reminder email a day before the actual experiment. Unlike those in the other treatments however, it contained a link to the exit-questionnaire. Access to the exit-questionnaire was password-protected. The email stated that they would have to keep the email with the link and they would receive the password during the experiment. The computer-voice interface announced and repeated this password after subjects had reported the result for the last coin toss.

Appendix F: Procedures and instructions for Experiment 2

The initial steps of Experiment 2 were essentially identical to treatment FORM in Experiment 1 (see Figure E.1). The only differences were that when subjects signed up in step 1, one had to make an appointment for Part B while Part A had to be completed at a pre-determined day, a week after the invitation. Moreover, the welcome stage (step 2) was on a web page for which subjects received an personalized link by email on the day when Part A had to be completed. Most importantly, after the socio-economic questionnaire which followed step 6, subjects received the instructions for the second coin tossing task (step 7). At the end of step 7, they were offered the choice whether they wanted to report the outcomes via Skype call or online form in the upcoming Part B. Finally, subjects entered their address to receive the payment by mail. This concluded Part A.

For Part B, those subjects selected for participation in Part B were contacted on the previously agreed date and time by an experimenter via Skype call.[8] They then had to report the results of their second set of coin tosses either via a Skype call or through an online form, depending on their choices in Part A. We used the same reporting protocol as in treatment FORM and CALL of Experiment 1, respectively.

The next pages show the instructions for the coin tossing task and choice of reporting channel in Experiment 2. As for Experiment 1, these are translations from German and content in frames was displayed in the subjects' web browser.

[8] Subjects who were not randomly selected to participate in Part B received an email notification on the day following Part A that informed them about the cancellation.

Step 7 (i) in Experiment 2: Announcement

Preparation for Part B:

In Part B, you again have the opportunity to earn up to CHF 20. We again ask you to **toss the coin ten times** *and to write down the results on paper.* **You will not report the results until Part B of the study.**

Please note that only Part A or Part B will be paid out. This will be determined at random at the end of the study.

Please have the coin, the paper, and the pencil ready. Using a second payment table, you can see whether you won the toss or not. Each win increases **your income by CHF 2.-**, *meaning that you can earn up to CHF 20.-*

Here is another example of a payment table :

		
Yield	CHF 2.-	CHF 0.-

You win if you toss tails, and your income is increased by CHF 2. -
If you toss heads, you lose and will not earn anything more.

Please click on "continue" to continue.

[continue-button]

Step 7 (ii) in Experiment 2: Instructions for the second coin tossing task

Please pick up the coin and toss it 10 times. **Please note the corresponding result for each toss, i.e., heads or tails.** *The payment table below shows for every toss the result that will allow you to earn CHF 2.-*

You may begin tossing the coin.

Toss 1

		
Yield	CHF 0.-	CHF 2.-

...

[The corresponding pictures for tosses 2 to 9 are omitted. The sequence was "Heads" /"Tails" /"Tails" /"Heads" /"Heads" /"Heads" /"Heads" /"Heads".]

Toss 10

		
Yield	CHF 0.-	CHF 2.-

Please click on "continue" when you have noted all 10 tosses.

[continue-button]

Step 7 (iii) in Experiment 2: Choice of the communication channel to report the results from the second coin tossing task

We now ask you to inform us **how you wish to report the results of the coin toss in Part B**. *You have two options (Your selection will not influence whether you will be chosen to participate in Part B):*

Option 1: Online form. *You will receive a link to the online form at the end of the Skype call (without video). The reporting of the results of the coin toss takes place in the same way as in PPart A.*

Option 2: Orally by Skype. *Following the Skype call, the study conductor will ask you to report the results of the coin tosses orally by Skype (without video), i.e., you must either say "heads" or "tails".*

Please note that no further questions will be asked in either option. Now decide between Option A and Option B:

[radio-button] With the online form

[radio-button] Orally by Skype

[continue-button]

The order in which the options were presented was randomized and counterbalanced across subjects.

References

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. Journal of Human Resources, 8(4), 436–455.

Oaxaca, Ronald (1973). Male-Female Wage Differentials in Urban Labor Markets. International Economic Review, 14(3), 693–709.