For Online Publication

Honesty in the Digital Age

Online Appendix

Michel André Maréchal ⓡ Alain Cohn ⓡ Tobias Gesche[1]

Contents:

[1] ⓡ indicates randomized order of authorship (Key for AEA Random Author Order Archive: hWk7nHk5ChCm). Maréchal: michel.marechal@econ.uzh.ch, Department of Economics, University of Zurich; Cohn: adcohn@umich.edu, School of Information, University of Michigan; Gesche: tgesche@ethz.ch, Center for Law & Economics, ETH 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 =1 if:	$y_i \in \{8, 9, 10\}$	$y_i \in \{7, 8, 9, 10\}$	$y_i \in \{9, 10\}$
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 =1 if:	$y_i \in \{6, 7\}$	$y_i \in \{6\}$	$y_i \in \{6, 7, 8\}$
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 =1 if:	$y_i \in \{8, 9, 10\}$	$y_i \in \{7, 8, 9, 10\}$	$y_i \in \{9, 10\}$
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 =1 if:	$y_i \in \{6, 7\}$	$y_i \in \{6\}$	$y_i \in \{6, 7, 8\}$
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 =1 if:	$y_i \in \{8, 9, 10\}$	$y_i \in \{7, 8, 9, 10\}$	$y_i \in \{9, 10\}$
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 =1 if:	$y_i \in \{6, 7\}$	$y_i \in \{6\}$	$y_i \in \{6, 7, 8\}$
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***	0.080***	0.081***
	(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	
		(0.024)	
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.

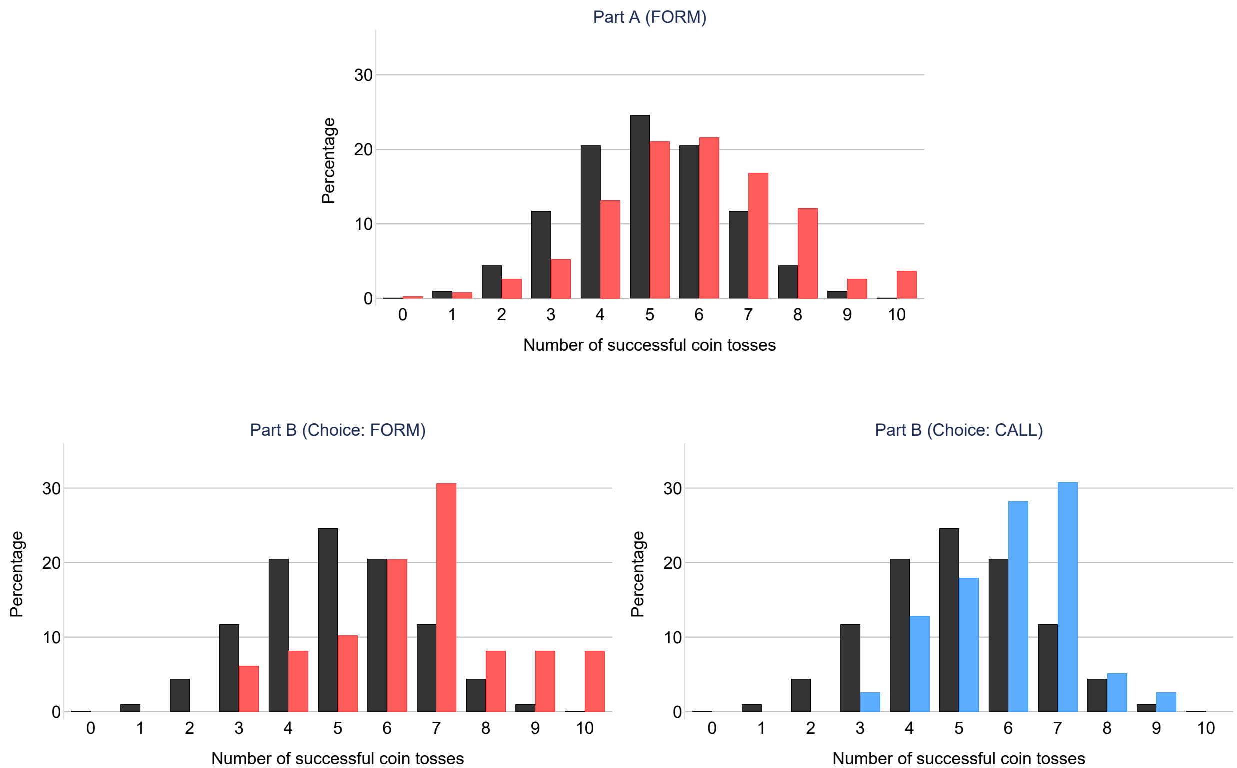
In the first wave of the main experiment we included a measure of social distance (the "Inclusion of Other in the Self" scale, see Aron et al., 1992). Subjects had to indicate how close they felt to the experimenter by selecting one out of five pairs of circles that varied by how much they overlap (which 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

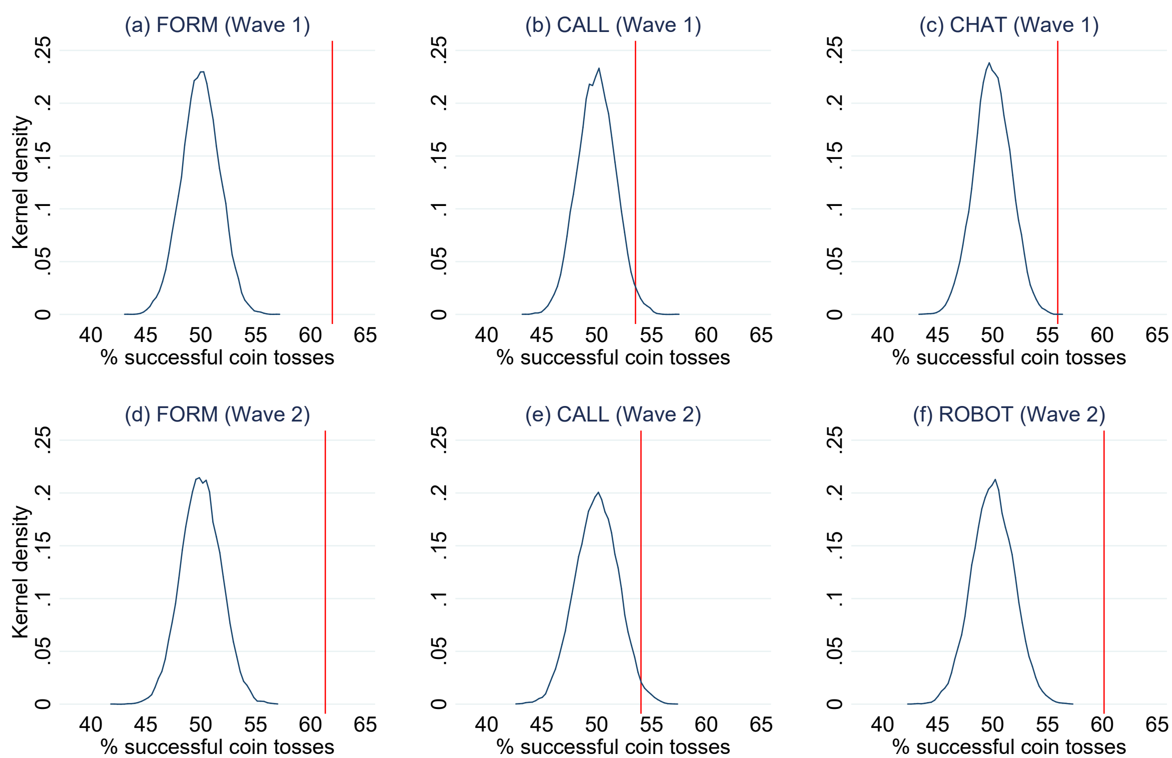


Notes: Colored bars depict actual observations by reporting channel; blue=choice to report via call, red=choice to report via form, and black bars depict the distribution expected under truthful reporting.

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 $p<0.0001$ (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.

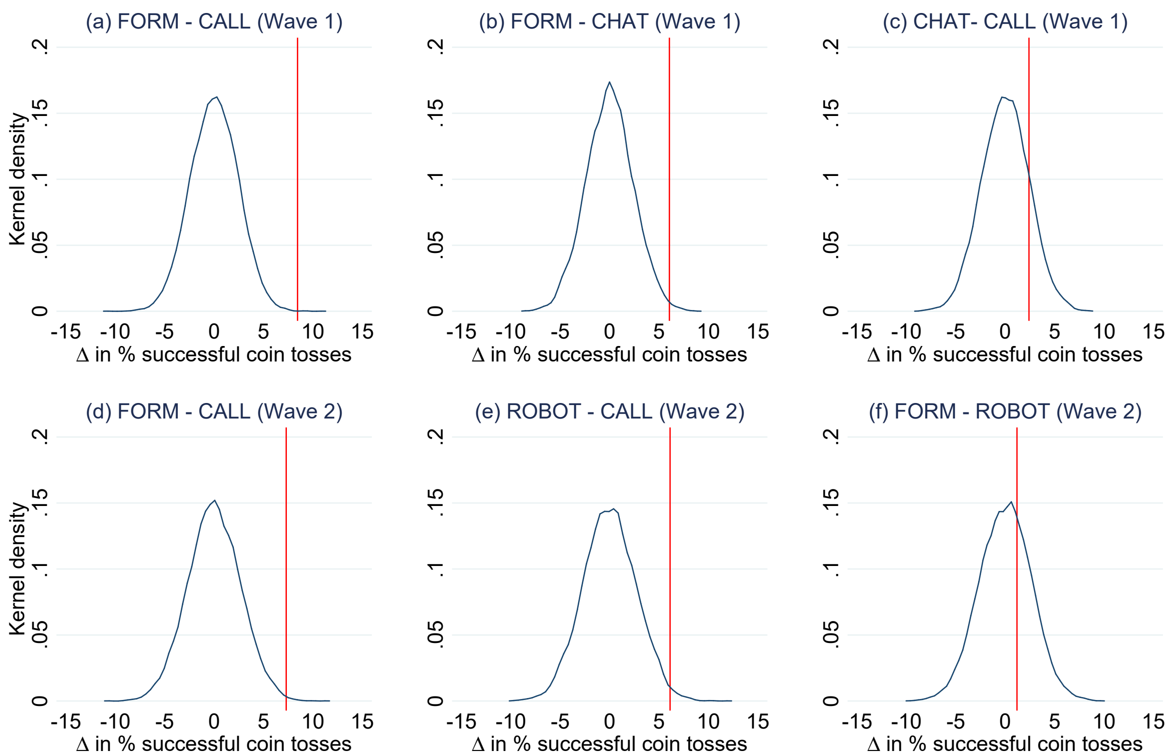
We next focus on treatment differences, and ask whether they could have occurred simply by chance. Figure C.2 presents the distributions of simulated treatment effects (assuming that all subjects reported truthfully) and the actual treatment effects represented by the red vertical lines. The probability 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 ($p<0.0001$ 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 ($p<0.0001$). 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.

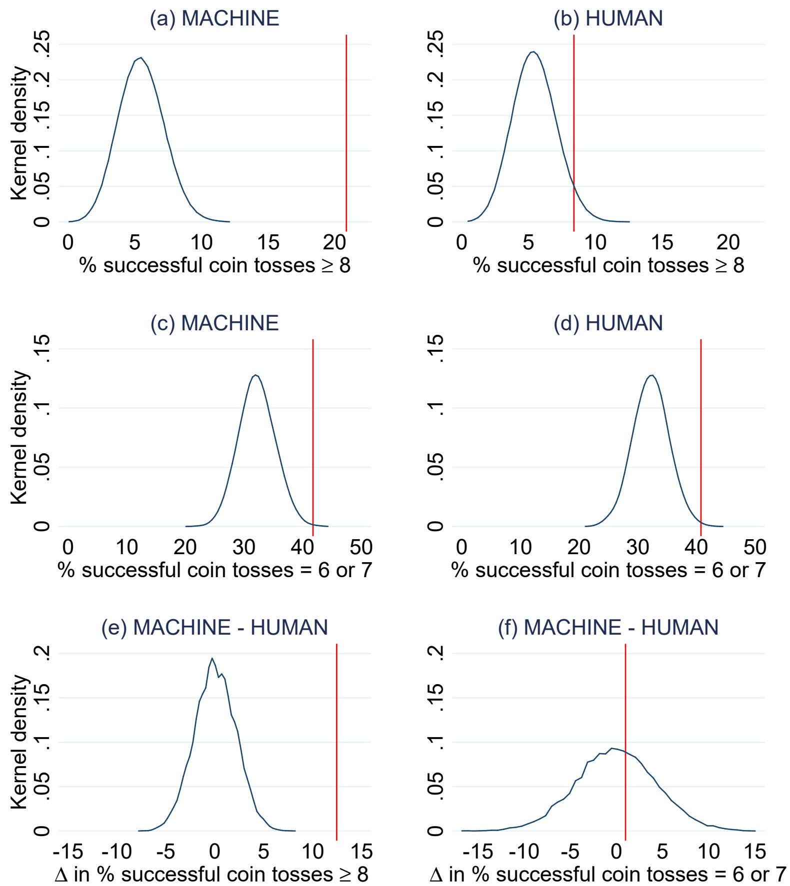
[3]We report two-sided simulated p-values for all tests that compare differences across treatments.

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



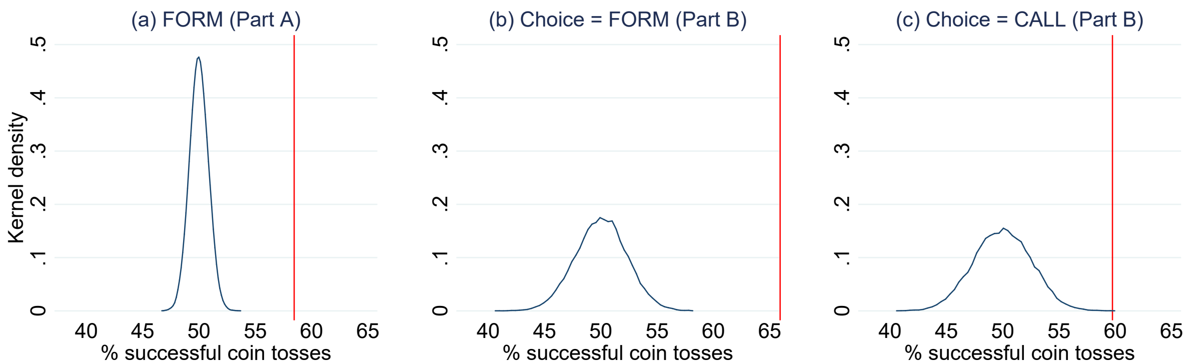
Notes: Panels (a) to (f) display kernel densities for the differences in successful coin flips between treatments for each of the 10,000 simulated samples, assuming n$\times 10$ truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the 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)



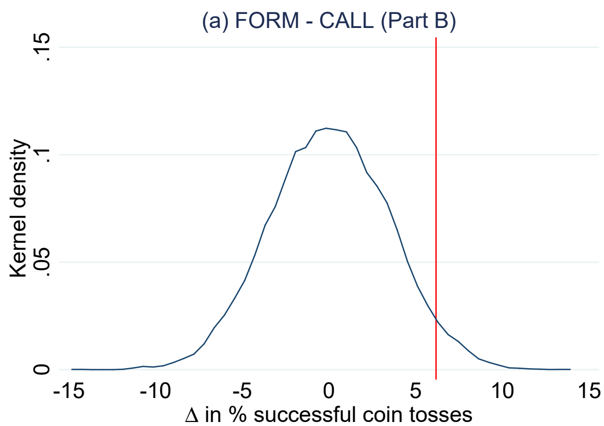
Notes: Panels (a) to (f) display kernel densities for the differences in successful coin flips between treatments for each of the 10,000 simulated samples, assuming $n \times 10$ truthfully reported coin tosses (i.e., a success probability of 50%), where n corresponds to the actual sample size of the 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 ($p<0.0001$ 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). Appendix G features detailed instructions for the survey experiment.

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.

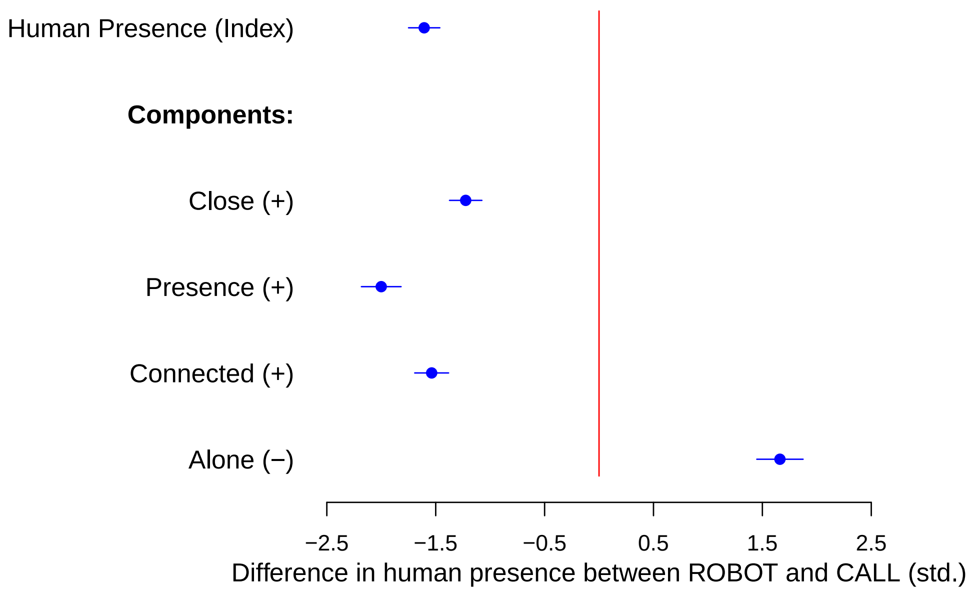
Results: We standardized the responses for each question 1 to 4 using the mean and standard deviation in the ROBOT condition. We then created a human presence index using the unweighted average of the standardized responses.[6] We followed the same procedure to construct an index of social image concerns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

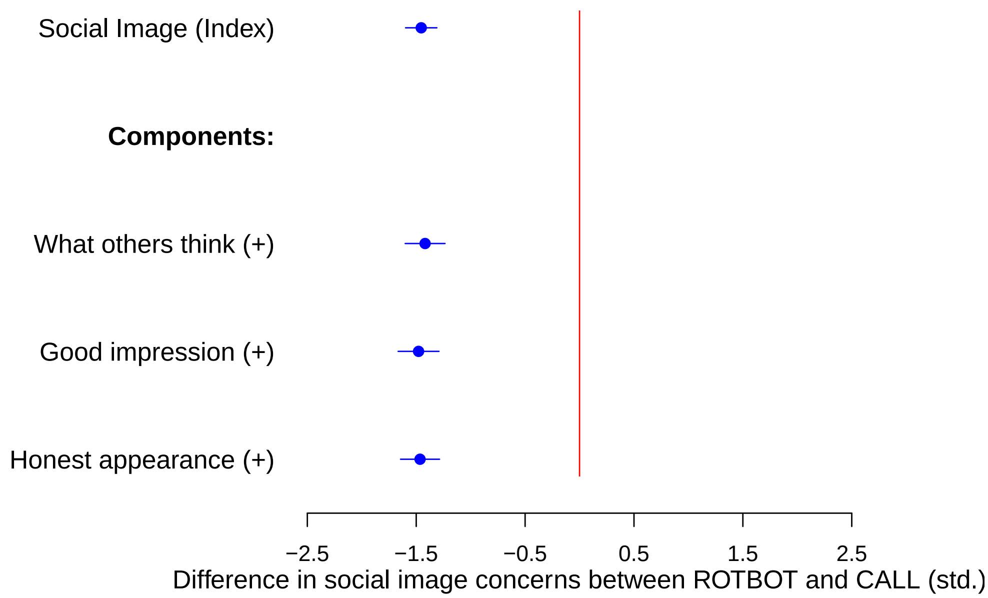
Figure D.2 shows that subjects were substantially less concerned about their social image in ROBOT than in CALL. The score of the social image index is roughly 1.5 standard deviations lower in ROBOT. The results are again remarkably consistent across each individual component of the index.

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.

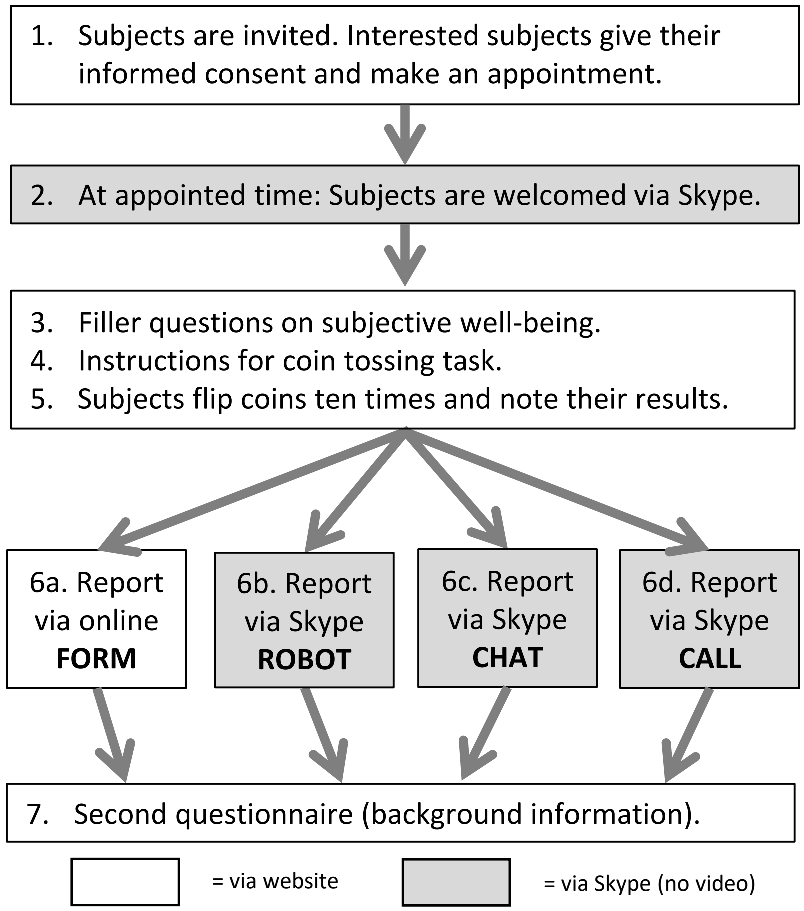
Appendix E: Procedures and instructions for Experiment 1

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.

Appendix G: Instructions for the survey experiment

The experiment had the following structure: After giving consent to participate and passing the attention check, subjects read a short description of their task (Step 1). In Step 2 subjects received (i) an outline of the main features of Experiment 1 and (ii) a more detailed description of treatment CALL and ROBOT. In Step 3 we instructed subjects to (i) take the perspective of a participant in the described experiment, (ii) answer four questions capturing their perception of human presence, and (iii) answer three questions about social image concerns. We used a within-subjects design, meaning that subjects first answered all questions for one treatment, and then for the other treatment (Step 4). The order of treatments CALL and ROBOT was randomized across subjects. The order of the questions within blocks (ii) and (iii) of Step 3 was also randomized but kept constant across treatments. Finally, in Step 5, participants predicted the average number of successful coin tosses reported in treatment ROBOT and CALL of Experiment 1.

The next pages contain the instructions for the survey experiment. The content in frames was displayed in subjects' web browser.

Description of the task

YOUR TASK

We recently conducted an experiment to get a better understanding of how people make decisions. Like any experiment, participants were randomly assigned to different conditions.

(As an analogy, think of a medical study where patients are randomly assigned to one of two conditions. In one condition, the patients receive a pill with an active ingredient while patients in the other condition receive a placebo pill.)

Your task is to **take the perspective of the participants** and answer some questions about how you would feel in the different conditions. Then, we will ask you to predict how participants behaved in each condition.

Additional bonus:

- Depending on the accuracy of your predictions, you can earn **additional money (up to $2).**
- We will also ask you a question about a specific design feature of the experiment. If you answer that question correctly, you will earn an **additional $0.5**
- It is therefore important that you read the information carefully!

Step 2 (i): Basic description of Experiment 1

The Experiment

The experiment was conducted remotely via the communication software Skype. A few days before the start of the experiment, participants had to indicate their preferred date and time because each participant participated individually. Participants were also asked to provide their Skype name and personal details (including their home address to receive their payments). On the agreed date and time each participant was contacted by a person who carried out the experiment (from now on we will refer to this person as the other person):

- The other person contacted participants via Skype's chat.
- Then, the other person sent participants a link to an online survey. Participant had to click on the link to open the survey in the web browser and then had to answer some questions about their life satisfaction.
- After that, participants performed a coin-tossing task. They had to take a coin, toss it 10 times and take note of each coin flip outcome (e.g., HEADS for coin flip 1, TAILS for coin flip 2 etc.).
- Before they started to toss the coin, participants were informed that they could earn money depending on the outcomes they reported at a later stage. Every time a participant reported their coin landing HEADS, they earned $2.
- Every time a participant reported their coin landing TAILS, they earned nothing.
- Thus, participants could earn up to $20 (10 x $2).

Step 2 (ii): Detailed description of treatment CALL and ROBOT in Experiment 1

Two experimental conditions

After the coin-tossing task, participants were randomly assigned to **one of two conditions**. These conditions differed in how participants had to report the outcomes of their coin tosses (see below for more details).

Condition CALL

- In this condition, participants had to **call back the other person** on Skype (without video) to report the outcomes of their coin tosses.
- The other person reminded the participants that every time they reported HEADS they would earn $2 (and $0 otherwise).
- For each coin toss, **the other person** asked participants to report the outcome and then participants had to answer by saying HEADS or TAILS. No further questions were asked and participants knew this.

Condition ROBOT

- In this condition, participants had to **call an interactive voice response system via Skype** (without video) to report the outcomes of their coin tosses. The voice response system used pre-recorded audio of the other person.
- Participants were reminded that every time they reported HEADS they would earn $2 (and $0 otherwise).
- For each coin toss the **voice response system** asked participants the outcome and then participants had to answer by saying HEADS or TAILS. No further questions were asked and participants knew this.

The study was completed after the reporting stage. A few days later, participants received their earnings from the coin-tossing task by mail.

Please make sure that you have carefully read the description of the experiment and that you understand the difference between the two conditions. If so, you can click on the CONTINUE button.

Step 3 (i): Taking the perspective of a participant

Taking the Perspective of the Participants

In the following, we will ask you some questions about **how you would feel as a participant** in the experiment. We will ask you those questions separately for each condition.

Step 3 (ii): Perception of human presence

Subjects were randomly assigned to treatment ROBOT or CALL. In the following we reproduce the instructions for treatment ROBOT.

Taking the Perspective of the Participants

Imagine you are a participant in the **ROBOT** condition (that is, you have to report the outcomes of the coin tosses by calling the voice recording system on Skype).

Note: If you do not remember the details of the **ROBOT** condition, click the Help-button at the bottom of this page.

As you report of the outcomes of you coin tosses to the voice response system (that uses the pre-recorded voice of the other person)...

How close would you feel to the other person?

Not at all 1	2	3	4	5	6	Very much 7
◯	◯	◯	◯	◯	◯	◯

How strongly would you feel the presence of the other person

Not at all 1	2	3	4	5	6	Very much 7
◯	◯	◯	◯	◯	◯	◯

How connected would you feel to the other person?

Not at all 1	2	3	4	5	6	Very much 7
◯	◯	◯	◯	◯	◯	◯

To what extent would you feel that you are alone?

Not at all 1	2	3	4	5	6	Very much 7
◯	◯	◯	◯	◯	◯	◯

Step 3 (iii): Social image concerns

The treatment was the same as in Step 3 (ii).

> Imagine you are a participant in the **ROBOT** condition (that is, you have to report the outcomes of the coin tosses by calling the other person on Skype).
>
> Note: If you do not remember the details of the **ROBOT** condition, click the Help-button at the bottom of this page.
>
> As you report of the outcomes of you coin tosses to the voice response system (that uses the pre-recorded voice of the other person)...
>
> How concerned would you be about what the other person thinks about you?
>
> | Not at all | | | | | | Very much |
> |---|---|---|---|---|---|---|
> | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
> | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |
>
> How much would you care about leaving a good impression on the other person?
>
> | Not at all | | | | | | Very much |
> |---|---|---|---|---|---|---|
> | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
> | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |
>
> How important would it be for you that the other person thinks you are honest
>
> | Not at all | | | | | | Very much |
> |---|---|---|---|---|---|---|
> | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
> | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |

Step 4 (i): Transition to the second treatment

> We will now ask you how participants experienced the **CALL** condition.

Step 4 (ii) and (iii): Repeat Step 3 (ii) and (iii) for the second treatment

The only difference was that the sentence "As you report of the outcomes of your coin tosses to the voice response system (that uses the prerecorded voice of the other person) ..." in ROBOT was replaced with "As you report the outcomes of your coin tosses to the other person on the Skype call ..." for CALL.

Step 5: Predicting behavior in Experiment 1

You now have the opportunity **to earn additional money**. In the experiment, participants had to toss a coin ten times. What do you think, how many times did the participants report that the coin landed HEADS (+$2), on average, in each condition? Try your best to be accurate, as the **most accurate 20% of participants will receive an additional payment of $1 for every estimate** (you can therefore earn up to 2 x $1 on this page).

Condition CALL (participants reported coin tosses by calling the other person on Skype)

On average, participants in CALL reported X times out of ten that the coin landed HEADS. *Enter a number between 0.0 and 10.0 for X, you can enter one decimal place.*

Condition ROBOT (participants reported coin tosses by calling an interactive voice response system on Skype)
On average, participants in ROBOT reported X times out of ten that the coin landed HEADS. *Enter a number between 0.0 and 10.0 for X, you can enter one decimal place*

Please click the button below if you entered your guesses.

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.