

"I'm just being honest."
When and why honesty enables help versus harm

Emma Levine & David Munguia Gomez
The University of Chicago, Booth School of Business

Supplemental Online Materials

Contents

1. Additional results and information from reported studies.....	1
1.1. Additional measures and results in Study 1.....	1
1.2. Additional results from Study 2.....	5
1.4. Preregistration information.....	11
2. Scenarios and Results from pilot study.....	12
3. Supplemental Studies.....	17
3.1. Study S1: The effects of honesty on selfishness.....	17
3.2. Study S2: Independently manipulating honesty and self-interest.....	25
3.3. Study S3: Manipulating whether omission is possible.....	32
3.4. Study S4: Manipulating whether omission is possible (negotiations).....	38
3.5. Study S5: Asymmetric views of selfish honesty: Actors versus targets.....	41
3.6. Study S6: Asymmetric views of selfish honesty: Actors versus targets versus observers.....	45

1. Additional results and information from reported studies

1.1. Additional measures and results in Study 1

In Study 1 of the main manuscript, we present only the statistics and planned contrasts corresponding to our key hypothesis tests. Specifically, we compared judgments among individuals who made the selfish choice in each condition (Control, Selfish Honesty, Altruistic Honesty). Below, we provide the results of all participants who completed the study.

We also report all measures we collected in this study. In addition to the measures we report in the main manuscript, participants rated the extent to which they felt the following emotions after making their decision: Guilty, Proud, Excited, Happy, Pleased, Ashamed, and Distressed (1 = “Not at all”; 7 = “Very”). Guilty, Ashamed, and Distressed loaded on a single factor and were combined to create a scale of *Negative Affect* ($\alpha = .89$). Proud, Excited, Happy, and Pleased loaded together on a separate factor and were combined to create a scale of *Positive Affect* ($\alpha = .88$). We report the results of these measures below.

Table S1. Choice x Decision judgments in Study 1

Panel A. Ethical Decision

Choice	Condition					
		Control	Selfish Honesty	Altruistic Honesty	Total	
	Inaction (Omission)	<i>M</i>	6.69	6.00	5.01	5.91
		<i>SD</i>	0.62	1.07	1.59	1.39
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	3.25	5.23	2.10	3.94
		<i>SD</i>	1.52	1.64	0.96	1.97
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	5.80	2.50	6.42	6.13
		<i>SD</i>	1.30	--	1.17	1.40
		<i>n</i>	5	1	18	24
	Total	<i>M</i>	5.01	5.35	4.12	4.82
		<i>SD</i>	2.05	1.59	2.14	2.00
		<i>n</i>	101	102	102	305
Main effect of Choice		$F(2, 296) = 90.36, p < .001, \eta_p^2 = .379$				
Main effect of Condition		$F(2, 296) = 3.92, p = .021, \eta_p^2 = .026$				
Choice x Condition		$F(4, 296) = 16.37, p < .001, \eta_p^2 = .181$				

Panel B. Honest Justification

Choice	Condition					
		Control	Selfish Honesty	Altruistic Honesty	Total	
	Inaction (Omission)	<i>M</i>	5.79	5.08	3.89	4.91
		<i>SD</i>	1.12	1.25	1.59	1.60
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	3.38	5.34	2.08	4.01
		<i>SD</i>	1.52	1.62	1.02	2.00
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	3.80	3.00	6.39	5.71
		<i>SD</i>	2.02	--	0.93	1.68
		<i>n</i>	5	1	18	24
	Total	<i>M</i>	4.54	5.26	3.62	4.48
		<i>SD</i>	1.81	1.56	1.99	1.91
		<i>n</i>	101	102	102	305
Main effect of Choice		$F(2, 296) = 25.92, p < .001, \eta_p^2 = .149$				
Main effect of Condition		$F(2, 296) = 0.45, p = .64, \eta_p^2 = .003$				
Choice x Condition		$F(4, 296) = 19.86, p < .001, \eta_p^2 = .212$				

Panel C. Selfish Justification

Choice	Condition					
		Control	Selfish Honesty	Altruistic Honesty	Total	
	Inaction (Omission)	<i>M</i>	3.43	3.56	4.82	4.00
		<i>SD</i>	1.10	0.91	1.20	1.29
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	6.29	5.28	6.20	5.78
		<i>SD</i>	0.79	1.33	0.97	1.22
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	2.20	1.67	3.07	2.83
		<i>SD</i>	0.84	--	1.26	1.22
		<i>n</i>	5	1	18	24
	Total	<i>M</i>	4.73	4.93	5.05	4.90
		<i>SD</i>	1.79	1.46	1.58	1.61
		<i>n</i>	101	102	102	305
Main effect of Choice		$F(2, 296) = 114.13, p < .001, \eta_p^2 = .435$				
Main effect of Condition		$F(2, 296) = 8.40, p < .001, \eta_p^2 = .054$				
Choice x Condition		$F(4, 296) = 5.34, p < .001, \eta_p^2 = .067$				

Panel D. Moral Identity

		Condition				
		Control	Selfish Honesty	Altruistic Honesty	Total	
Choice	Inaction (Omission)	<i>M</i>	6.13	5.76	5.35	5.76
		<i>SD</i>	0.76	0.98	1.04	0.98
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	5.16	5.54	4.46	5.18
		<i>SD</i>	1.02	1.00	1.17	1.13
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	6.00	4.75	6.06	5.99
		<i>SD</i>	0.47	--	1.07	0.98
		<i>n</i>	5	1	18	24
	Total	<i>M</i>	5.66	5.57	5.13	5.45
		<i>SD</i>	1.00	0.99	1.24	1.11
		<i>n</i>	101	102	102	305
Main effect of Choice		$F(2, 296) = 14.47, p < .001, \eta_p^2 = .089$				
Main effect of Condition		$F(2, 296) = 2.94, p = .054, \eta_p^2 = .020$				
Choice x Condition		$F(4, 296) = 3.00, p = .019, \eta_p^2 = .039$				

Panel E. Positive Emotion

		Condition				
		Control	Selfish Honesty	Altruistic Honesty	Total	
Choice	Inaction (Omission)	<i>M</i>	4.47	3.76	3.34	3.90
		<i>SD</i>	1.73	1.55	1.51	1.68
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	3.18	3.33	3.06	3.22
		<i>SD</i>	1.51	1.64	1.31	1.53
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	4.90	5.50	3.74	4.05
		<i>SD</i>	1.21	--	1.78	1.71
		<i>n</i>	5	1	18	24
Total	<i>M</i>	3.88	3.43	3.30	3.54	
	<i>SD</i>	1.73	1.63	1.49	1.63	
	<i>n</i>	101	102	102	305	
Main effect of Choice		$F(2, 296) = 7.55, p = .001, \eta_p^2 = .049$				
Main effect of Condition		$F(2, 296) = 3.93, p = .021, \eta_p^2 = .026$				
Choice x Condition		$F(4, 296) = 1.58, p = 0.17, \eta_p^2 = .021$				

Panel F. Negative Emotion

		Condition				
		Control	Selfish Honesty	Altruistic Honesty	Total	
Choice	Inaction (Omission)	<i>M</i>	1.20	1.32	2.11	1.58
		<i>SD</i>	0.66	0.56	1.29	1.04
		<i>n</i>	48	19	44	111
	Option A (Selfish)	<i>M</i>	2.31	1.87	3.29	2.33
		<i>SD</i>	1.39	1.25	1.46	1.45
		<i>n</i>	48	82	40	170
	Option B (Altruistic)	<i>M</i>	1.00	4.00	1.33	1.38
		<i>SD</i>	0.00	--	0.75	0.86
		<i>n</i>	5	1	18	24
Total	<i>M</i>	1.72	1.79	2.43	1.98	
	<i>SD</i>	1.20	1.19	1.48	1.33	
	<i>n</i>	101	102	102	305	
Main effect of Choice		$F(2, 296) = 19.08, p < .001, \eta_p^2 = .114$				
Main effect of Condition		$F(2, 296) = 5.61, p = .004, \eta_p^2 = .036$				
Choice x Condition		$F(4, 296) = 3.37, p = .01, \eta_p^2 = .044$				

1.2. Additional results from Study 2

Below we report the same set of analyses as in the manuscript, split by recruitment location (Coffee shop versus MTurk) and levels of Honesty-Humility (low versus high). We also report additional analyses that include Honesty-Humility (HH) scores as a predictor of selective honesty.

Recruitment location. We recruited participants in two locations, a local coffee shop, known for fair trade practices and MTurk. As reported in the manuscript, we used different incentives across the coffee shop and MTurk samples, but other than this difference, the studies were identical across samples.

Overall, we found a similar pattern of results across our samples, with two exceptions: (1) the coffee shop sample did not significantly differ in their levels of honesty in the *Selfish Honesty* and *Altruistic Honesty* condition and (2) the *Altruistic Honesty* condition did not significantly affect the MTurk sample's distribution of choices.

Coffee shop sample (n = 206). We found that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 20.57, p < .001$). The coefficients of the multinomial logistic regression reveal that relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation ($b = 1.98, 95\% \text{ CI} = [1.12, 2.84]$). We also found that participants' distribution of choices was significantly different in the *Altruistic Honesty* condition and the *Control* condition ($\chi^2 = 13.81, p = .001$). Relative to the *Control* condition, participants in the *Altruistic Honesty* condition shifted their choices away from inaction and towards the altruistic allocation ($b = 1.49, 95\% \text{ CI} = [0.69, 2.29]$).

Though having the opportunity to tell a selfish truth (*Selfish Honesty* condition) had a qualitatively larger effect on participants' distribution of choices than having the opportunity to tell an altruistic truth (*Altruistic Honesty* condition) did, the effect of *Selfish Honesty* on selfish choice did not significantly differ from the effect of *Altruistic Honesty* on altruistic choice ($\chi^2 = 0.77, p = .380$). When we examine selective honesty, we found that participants were not significantly more likely to be honest in the *Selfish Honesty* condition than in the *Altruistic Honesty* condition ($\chi^2 = 0.0003, p = .987$); this pattern of results is different than for the overall sample.

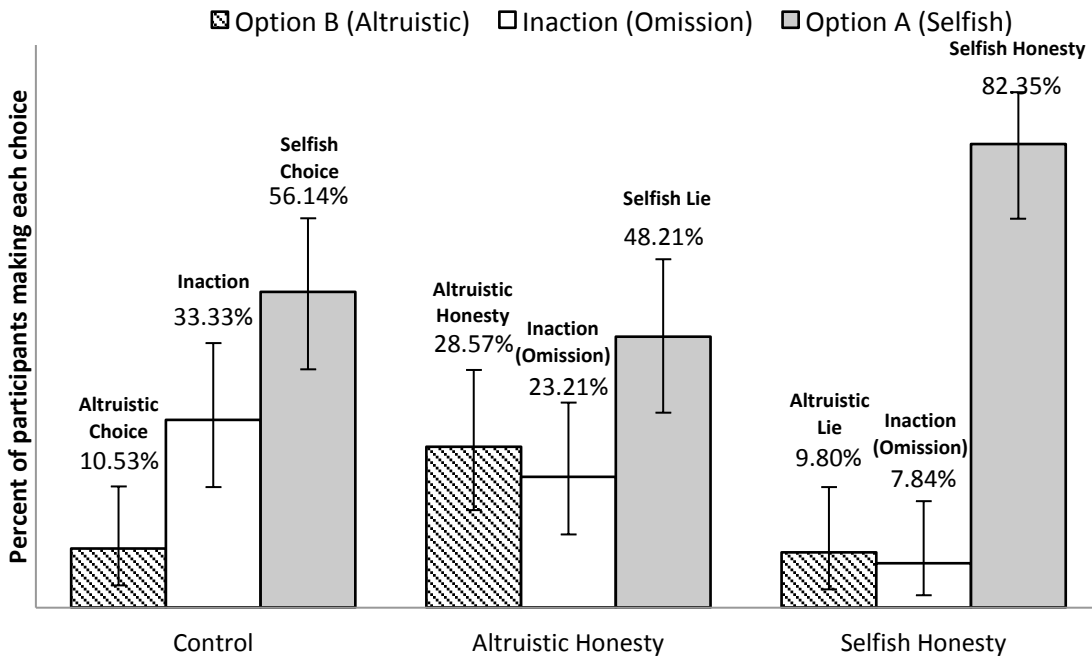
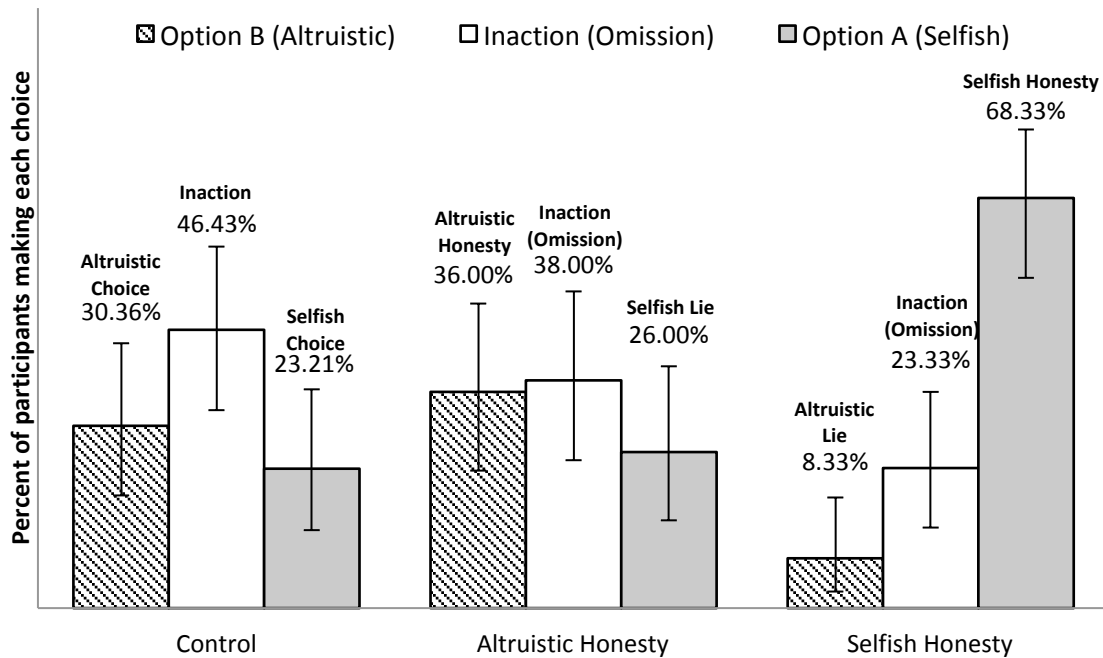
MTurk (n = 229). We found that that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 14.53, p = .001$). The coefficients of the multinomial logistic regression reveal that relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation ($b = 1.20, 95\% \text{ CI} = [0.41, 2.00]$). We do not find that participants' distribution of choices was significantly different in the *Altruistic Honesty* condition and the *Control* condition ($\chi^2 = 1.61, p = .448$); this pattern of results is different than for the overall sample. Though having the opportunity to tell a selfish truth (*Selfish Honesty* condition) had a qualitatively larger effect on participants' distribution of choices than having the opportunity to tell an altruistic truth (*Altruistic Honesty* condition) did, the effect of *Selfish Honesty* on selfish

choice did not significantly differ from the effect of *Altruistic Honesty* on altruistic choice ($\chi^2 = 1.65, p = .20$). When we examine selective honesty, we found that participants were significantly more likely to be honest in the *Selfish Honesty* condition than in the *Altruistic Honesty* condition ($\chi^2 = 45.77, p < .001$).

Honesty-Humility. As indicated in Footnote 4, participants indicated their agreement with the 10 Honesty-Humility scale items from the HEXACO-60 personality inventory (Ashton & Lee, 2009), after making their decisions. The Honesty-Humility scale reflects traits such as sincerity, fairness, unpretentiousness, and lack of greed (Ashton & Lee, 2009). Participants indicated their agreement with statements such as, “I would never accept a bribe, even if it were very large,” “Having a lot of money is not especially important to me,” and “I wouldn’t use flattery to get a raise or promotion at work, even if I thought it would succeed.” All items were anchored at 1 = “Strongly disagree” and 5 = “Strongly agree.” We calculated each participant’s average score for these items (using the scoring guideline in Ashton & Lee, 2009) to examine how their Honesty-Humility levels relate to their choices. We added these items after we had begun collecting data, so we do not have data on these items for 105 participants (out of the total 435). We started collecting these items after realizing that the rates of honesty and altruism were higher for participants recruited at the coffee shop than for the MTurk participants recruited in previous studies. We began collecting Honesty-Humility measures to see if differences in this personality trait could explain the observed differences in choice.

To examine how participants’ Honesty-Humility level was associated with their choices, we split the sample into participants with Honesty-Humility scores below and above the median (*Median* = 3.60, *Mean* = 3.50, *SD* = .72). The distribution of Honesty-Humility levels was slightly skewed (-.34), such that the mean of the distribution was slightly less than the median, and kurtosis was approximately normal (3.09). Figure S1 depicts the exact pattern of choices across conditions for participants with low ($n = 164$) and high ($n = 166$) Honesty-Humility scores.

Below, we report the same set of analyses reported in the manuscript, split by low and high Honesty-Humility scores, and we depict the distribution of choices in Figure S1. Overall, we found that the extent to which participants chose the selfish and altruistic options varied between low- and high-Honesty-Humility samples: Low Honesty-Humility participants were more likely than High Honesty-Humility participants to select the selfish option across conditions.

Figure S1. Effects of honesty on selfishness and altruism, by Honesty-Humility (Study 2)Panel A: Low Honesty-Humility ($N = 164$)Panel B: High Honesty-Humility ($N = 166$)

Note. Error bars reflect 95% confidence intervals. Low and high levels of Honesty-Humility are defined relative to the median.

Low Honesty-Humility (n = 164). We found that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 9.44, p = .009$). The coefficients of the multinomial logistic regression reveal that relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation, as revealed by the coefficients of the multinomial logistic regression ($b = 1.83, 95\% \text{ CI} = [0.66, 3.00]$). We also found that participants' distribution of choices was marginally significantly different in the *Altruistic Honesty* condition and the *Control* condition ($\chi^2 = 5.68, p = .058$). Relative to the *Control* condition, participants in the *Altruistic Honesty* condition shifted their choices away from inaction and towards the altruistic allocation ($b = 1.36, 95\% \text{ CI} = [0.19, 2.53]$). Though having the opportunity to tell a selfish truth (*Selfish Honesty* condition) had a qualitatively larger effect on participants' distribution of choice than having the opportunity to tell an altruistic truth (*Altruistic Honesty* condition) did, the effect of *Selfish Honesty* on selfish choice was not significantly different than the effect of *Altruistic Honesty* on altruistic choices ($\chi^2 = 0.36, p = .548$). When we examined selective honesty, we found that participants were significantly more likely to be honest in the *Selfish Honesty* condition than in the *Altruistic Honesty* condition ($\chi^2 = 31.10, p < .001$).

High Honesty-Humility (n = 166). We found that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 22.25, p < .001$). The coefficients of the multinomial logistic regression reveal that relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation, as revealed by the coefficients of the multinomial logistic regression ($b = 1.77, 95\% \text{ CI} = [0.87, 2.67]$). We did not find that participants' distribution of choices was significantly different in the *Altruistic Honesty* condition and the *Control* condition ($\chi^2 = 0.78, p = .678$). Having the opportunity to tell a selfish truth (*Selfish Honesty* condition) had a significantly larger effect on participants' distribution of choice than having the opportunity to tell an altruistic truth (*Altruistic Honesty* condition) did ($\chi^2 = 5.74, p = .017$). When we examined selective honesty, we found that participants were significantly more likely to be honest in the *Selfish Honesty* condition than in the *Altruistic Honesty* condition ($\chi^2 = 11.47, p = .001$).

Coffee shop and MTurk levels of Honesty-Humility. The levels of Honesty-Humility for the Coffee shop sample ($n = 101, \text{Median} = 3.7, \text{Mean} = 3.68, \text{SD} = .61$) was higher than for the MTurk sample ($n = 229, \text{Median} = 3.5, \text{Mean} = 3.42, \text{SD} = .75$), suggesting that many of the sample differences may have been driven by Honesty-Humility differences.

Multinomial logistic regression on choice with Honesty-Humility as a factor. In Table S2 below, we depict the results of a multinomial logistic regression on choices, including the Honesty-Humility score (z-scored) and its interactions as factors. As in the regression reported in the manuscript (which omitted Honesty-Humility), we found that *Selfish Honesty* ($\chi^2 = 26.84, p < .001$) and *Altruistic Honesty* ($\chi^2 = 5.46, p = .065$) each (marginally) significantly influenced participants' distribution of choices. The *Selfish Honesty* and *Altruistic Honesty* conditions primarily influenced choice by shifting participants away from inaction and towards the selfish allocation ($b = 1.75, 95\% \text{ CI} = [1.02, 2.48]$) and the altruistic allocation ($b = 0.87, 95\% \text{ CI} = [0.13, 1.61]$), respectively. We found a main effect of Honesty-Humility on selfish allocation

choice, such that participants high in Honesty-Humility were less likely to choose the selfish allocation ($b = -0.70$, 95% CI = [-1.19, -0.22]).

Table S2. Multinomial logistic regression on choices, including Honesty-Humility scores (Study 2)

Choice	Condition	b	SE	p	95% CI	
					Lower	Upper
Altruistic allocation	Altruistic Honesty	0.868	0.379	0.022	0.125	1.612
	Selfish Honesty	0.345	0.505	0.494	-0.644	1.335
	Honesty-Humility (HH)	0.301	0.295	0.307	-0.277	0.880
	Altruistic Honesty x HH	-0.480	0.428	0.262	-1.318	0.358
	Selfish Honesty x HH	-0.810	0.487	0.096	-1.764	0.145
	Constant	-0.770	0.280	0.006	-1.319	-0.221
Selfish allocation	Altruistic Honesty	0.208	0.341	0.542	-0.461	0.877
	Selfish Honesty	1.751	0.371	0.000	1.023	2.479
	Honesty-Humility (HH)	-0.704	0.246	0.004	-1.186	-0.222
	Altruistic Honesty x HH	-0.350	0.411	0.395	-1.156	0.456
	Selfish Honesty x HH	0.071	0.371	0.847	-0.655	0.798
	Constant	-0.063	0.223	0.776	-0.500	0.373

Post-hoc tests

Overall effect of Altruistic Honesty

$\chi^2 = 5.46$, $p = .065$

Overall effect of Selfish Honesty

$\chi^2 = 26.84$, $p < .001$

Effect of Altruistic Honesty on Altruism vs. Effect of Selfish Honesty on Selfishness

$\chi^2 = 3.32$, $p = .069$

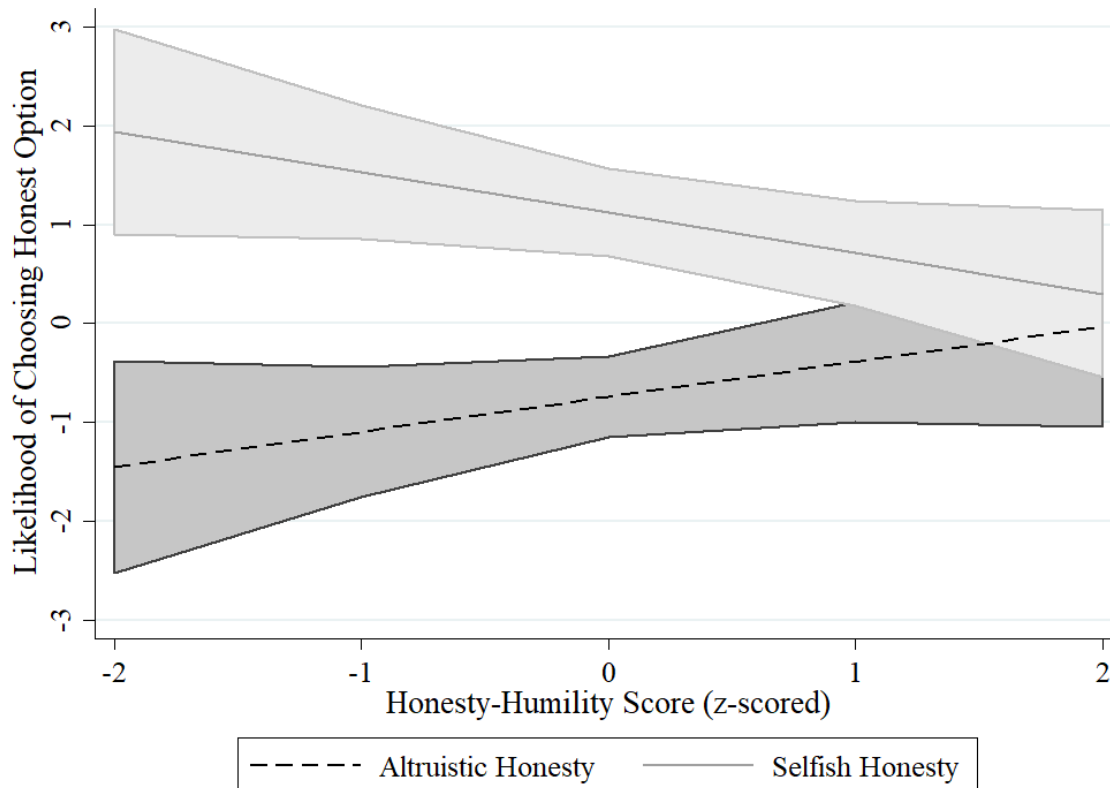
Note. Altruistic allocation was coded as 1, Inaction (omission) was coded as 2, and Selfish allocation was coded as 3. Inaction (omission) is the base outcome. The Honesty-Humility score was z-scored (mean = 3.50).

Firth logistic regression on honest choice with Honesty-Humility as a factor. To examine whether individual differences in Honesty-Humility moderate selective honesty, we depict the results of a penalized maximum likelihood logistic regression (using the *firthlogit* command in Stata) on Honest choice (1 = honest choice, 0 = all other choices), where we include Condition (1 = *Selfish Honesty*, 0 = *Altruistic Honesty*), Honesty-Humility score (z-scored), and the Condition x Honesty-Humility interaction as independent variables.

We found a main effect of *Condition* ($b = 1.86$, $SE = 0.31$, $p < .001$, 95% CI = [1.26, 2.47]), such that participants were more likely to choose the honest option when honesty was selfish rather than altruistic. There was no significant main effect of Honesty-Humility scores ($b = 0.35$, $SE = 0.25$, $p = .148$, 95% CI = [-0.13, 0.84]). There was, however, a significant Honesty

Condition x Honesty-Humility score interaction ($b = -0.77$, $SE = 0.33$, $p = .019$, 95% CI = [-1.40, -0.13]), such that participants who were low in Honesty-Humility were more likely to exhibit selective honesty (more likely to be honest when honesty is selfish than altruistic) than were participants who were high in Honesty-Humility. We plot the relationship between choosing the honest option and honest-humility when honesty is selfish versus altruistic in Figure S2 below.

Figure S2. Selective honesty, based on Honesty-Humility scores (Study 2)



Note. The dependent variable corresponds to the likelihood of selecting the honest option (as opposed to the dishonest or omission options), based on a penalized maximum likelihood logistic regression. The plot depicts the predicted likelihood of honesty, across levels of Honesty-Humility (z-scored) and whether honesty was altruistic or selfish. The shaded errors bars reflect 95% confidence intervals.

1.4. Preregistration information

Study 2: <https://aspredicted.org/8th62.pdf> and <https://aspredicted.org/7zn8b.pdf>

We preregistered two sets of analyses for Study 2. As indicated in the manuscript and in SOM 1.2, we preregistered the second set of analyses after adding the Honesty-Humility measure.

We also wish to note how the reported analyses differ from the preregistered ones:

- First preregistration (first link): The preregistration did not specify the precise analysis that we would use to examine “whether the frequency with which people choose the selfish allocation, the altruistic allocation, and the omission differs between each of the three cells”. We used a multinomial logistic regression.
- Amendment (second link): We preregistered a logistic regression, but conducted a penalized maximum likelihood logistic regression. As reported in the manuscript, this latter regression maintains the assumptions of logistic regression, but corrects for rare events (Firth 1993).
- Amendment (second link): First, the pre-registered regression did not take into account that predicting honesty as the dependent variable necessarily excluded the Control condition (because it has no honest option), which led us to incorrectly specify its list of factors. The present regression corrects that oversight by correctly specifying the factors.

Study 3: <https://aspredicted.org/jp783.pdf>

We wish to note how the reported analyses differ from the preregistered ones:

- To examine whether honesty enables self-interest versus harm we preregistered two sets of analyses: (1) a tabulation of choice for each Decision context (e.g., *Selfish/Harmful Honesty*) in each Consequence condition (e.g., *Self only*), followed by planned contrasts, and (2) a logistic regression. In the manuscript we replace the tabulation of choice with a multinomial logistic regression. (The key difference between these analyses is that the multinomial logistic regression compares the distribution of choices between conditions using Wald tests instead of chi-square tests.)

Study 5: <https://aspredicted.org/m6p4s.pdf>

The reported analyses do not differ from the preregistered ones.

2. Scenarios and Results from pilot study

We ran a pilot study with the goal of collecting some descriptive data on how frequently people use and encounter others using honesty as a justification for harm, how moral they perceive this justification to be, and how using honesty as a justification compares to other possible ethical justifications. Participants read three scenarios in which a target used a moral principle as a justification for harming others. Within each scenario, participants were randomly assigned to read about and judge one moral justification (Honesty, fairness, rule-following, deservingness, or reciprocity). All materials and results are presented below.

Scenario 1

Alex and Thomas are colleagues in a marketing department and they are competing for a promotion. Alex has an opportunity to undermine Thomas. Specifically, Alex's boss asks him why the supply closet is always out of pens, and Alex happens to know that Thomas recently stole office supplies from the company.

Honesty condition:

Alex decides to tell Thomas's boss about the theft. When asked why he decided to tell Thomas's boss, Alex says, "I just wanted to be honest."

Fairness condition:

Alex decides to tell Thomas's boss about the theft. When asked why he decided to tell Thomas's boss, Alex said, "I just wanted the promotion decision to be fair."

Authority/rule-following condition:

Alex decides to tell Thomas's boss about the theft. When asked why he decided to tell Thomas's boss, Alex said, "I was just following the rules."

Deservingness condition:

Alex decides to tell Thomas's boss about the theft. When asked why he decided to tell Thomas's boss, Alex said, "Thomas deserves to be punished."

Reciprocity condition:

Alex decides to tell Thomas's boss about the theft. When asked why he decided to tell Thomas's boss, Alex said, "Thomas would have told my boss if I had done that."

Scenario 2

Alex and Thomas are colleagues in a marketing department and the two do not like each other. In the past, Thomas has told others in the office that Alex has a poor work ethic. During a staff meeting, Alex has an opportunity to undermine Thomas. Specifically, Alex is asked to share how the firm's client base is doing, and happens to know that Thomas recently lost a big client because he was negligent.

Honesty condition:

Alex decides to announce the details of this loss to the group. When asked why he decided to do this, Alex said, "I just wanted to be honest."

Fairness condition:

Alex decides to announce the details of this loss to the group. When asked why he decided to do this, Alex said, "I just wanted Thomas to be fairly judged according to his merit."

Authority/rule-following condition:

Alex decides to announce the details of this loss to the group. When asked why he decided to do this, Alex said, "My boss told me that it is my job to inform the group."

Deservingness condition:

Alex decides to announce the details of this loss to the group. When asked why he decided to do this, Alex said, "Thomas deserves to be punished."

Reciprocity condition:

Alex decides to announce the details of this loss to the group. When asked why he decided to do this, Alex said, "Thomas would have told the group too if I had done that."

Scenario 3

Alex and Thomas are siblings and they do not get along. In the past, Thomas told their parents that Alex cheated on a test. During a phone call with his parents, Alex has an opportunity to undermine Thomas. Specifically, Alex' parents ask how Thomas is doing and Alex happens to know that Thomas recently got a DUI.

Honesty condition:

Alex decides to tell his parents about his brother's DUI. When asked why he decided to do this, Alex said, "I just wanted to be honest."

Fairness condition:

Alex decides to tell his parents about his brother's DUI. When asked why he decided to do this, Alex says, "It is only fair that our parents know."

Rule-following condition:

Alex decides to tell his parents about his brother's DUI. When asked why he decided to do this, Peter says, "My parents demanded to know."

Deservingness condition:

Alex decides to tell his parents about his brother's DUI. When asked why he decided to do this, Alex said, "Thomas deserves it."

Reciprocity condition:

Alex decides to tell his parents about the DUI. When asked why he decided to do this, Alex said, "Thomas would have told them if it were me who got the DUI."

Dependent variables:

1. Can you think of a time in which someone used honesty (fairness, rule-following, deservingness, reciprocity) as a justification for harming or undermining another person like Alex did in the scenario above?

(1 = yes, 2 = no)

Please explain (open-ended)

2. **How often do you think most people do this** (use honesty [principle] as a justification for harming or undermining another person like Alex did in the scenario above)?

Multiple times/day (1) , daily (2), weekly (3), monthly (4), yearly (5), Rarely/never (6)

3. **Have you ever done something like this** (used honesty [principle] as a justification for harming or undermining another person like Alex did in the scenario above)?

(1 = yes, 2 = no)

Please explain (open-ended)

4. **How often do you do this** (use honesty [principle] as a justification for harming or undermining another person like Alex did in the scenario above)?

Multiple times/day (1) , daily (2), weekly (3), monthly (4), yearly (5), Rarely/never (6)

5. *Please rate Alex's **character** based on the scenario above: moral, trustworthy, ethical
(1 = not at all, 7 = very)

6. *Please rate Alex's **decision** [to tell Thomas' boss]: Justified, moral, acceptable, ethical
(1 = not at all, 7 = very)

*All items in questions 5 and 6 loaded together in a factor analysis (Varimax rotation, principal axis factoring) and thus, were combined into one measure of morality ($\alpha = .97$)

ResultsRecall:

		Can you think of a time in which someone used this principle as a justification?	Have you ever used this principle as a justification?
		(question 1)	(question 3)
		<u>% saying yes</u>	
TOTAL	Honesty	66.50%	28.60%
	Fairness	48.00%	27.30%
	Deservingness	63.20%	30.40%
	Reciprocity	59.90%	29.50%
	Rule-following	53.20%	26.90%
Scenario 1	Honesty	70.90%	35.20%
	Fairness	47.50%	28.80%
	Deservingness	71.90%	33.30%
	Reciprocity	50.00%	26.40%
	Rule-following	56.10%	29.60%
Scenario 2	Honesty	61.00%	15.50%
	Fairness	46.40%	12.50%
	Deservingness	51.00%	23.50%
	Reciprocity	58.60%	19.00%
	Rule-following	42.10%	19.60%
Scenario 3	Honesty	67.90%	35.70%
	Fairness	50.00%	40.40%
	Deservingness	65.50%	34.00%
	Reciprocity	70.70%	43.60%
	Rule-following	61.40%	31.60%

Frequency and morality:

How often do most people do this? (question 2: lower numbers mean more often)

	<u>Total</u>	<u>Scenario 1</u>	<u>Scenario 2</u>	<u>Scenario 3</u>
Honesty	3.82	3.80	4.02	3.64
Fairness	4.04	4.10	3.93	4.09
Deservingness	3.93	3.74	4.06	4.02
Reciprocity	3.93	4.15	4.03	3.60
Rule-following	3.97	3.91	3.94	4.05

How often do you do this? (question 4: lower numbers mean more often)

	<u>Total</u>	<u>Scenario 1</u>	<u>Scenario 2</u>	<u>Scenario 3</u>
Honesty	5.40	5.28	5.52	5.39
Fairness	5.48	5.41	5.63	5.40
Deservingness	5.40	5.33	5.24	5.62
Reciprocity	5.36	5.42	5.48	5.16
Rule-following	5.45	5.48	5.50	5.39

How moral is this? (questions 5 & 6: lower numbers mean less moral)

	<u>Total</u>	<u>Scenario 1</u>	<u>Scenario 2</u>	<u>Scenario 3</u>
Honesty	3.27	4.40	2.52	2.95
Fairness	3.49	3.92	2.86	3.65
Deservingness	3.24	3.94	2.99	2.74
Reciprocity	3.13	3.21	2.86	3.35
Rule-following	3.55	4.07	3.05	3.53

3. Supplemental Studies

3.1. Study S1: The effects of honesty on selfishness

Study S1 was identical to Study 1 in the main paper, with two exceptions. First, we did not include the *Altruistic Honesty* condition. Second, we used a laboratory rather than MTurk sample, and therefore, used higher monetary stakes.

Method

We preregistered this study on AsPredicted.org (<https://aspredicted.org/pe8ny.pdf>). However, the analyses (on choice) reported below differ from what we preregistered. We use multinomial logistic regression to analyze choice data, in order to be consistent with the approach to analyze Study 1. However, the preregistered analyses (chi-square tests) are available in the posted syntax files (see <https://osf.io/gz4xh/>) and the results conceptually replicate all results described below.

Participants. We aimed to recruit 200 adults from a university laboratory pool in Chicago, IL to participate in this study, based on the a priori goal of recruiting 100 participants per condition. We ended up with a final sample of 189 participants (47% female; $M_{\text{age}} = 30$ years, $SD = 13.62$) who completed the entire study.

Procedure and Materials. In this study, we randomly assigned participants to one of two conditions in a between-subjects design: *Control* vs. *Selfish Honesty*.

Participants played an economic game, nearly identical to the game used in Study 1 of the main manuscript. Participants were assigned to the role of Decider and paired with a confederate Receiver. As the Decider, participants had the opportunity to allocate money between themselves and their partner, the Receiver.

All participants started the game with a default allocation: \$2.50 for the Decider (the participant) and \$2.50 for the Receiver (the participant's confederate partner). This default allocation represents a fair, or equal, allocation. Participants, however, had the opportunity to change the allocation to one of the following options:

- Option A) \$4.50 for the Decider and \$0.50 for the Receiver (*selfish*)
- Option B) \$0.50 for the Decider and \$4.50 for the Receiver (*altruistic*)

In the *Control* condition, participants had three choices: they could do nothing and keep the default allocation (which we conceptualize as inaction), switch to Option A, or switch to Option B.

In the *Selfish Honesty* conditions, we told participants that the new allocation depended on their choice and the outcome of a random number generator that would generate a number between 1 and 9. Participants then had to report whether the number was ODD or EVEN. If participants reported that the number was ODD, the allocation would switch to Option A (*selfish*); if they reported that the number was EVEN, the allocation would switch to Option B (*altruistic*).

Participants could also choose not to report a number (i.e., to engage in omission/inaction), which would maintain the default allocation.

In the *Selfish Honesty* condition, the number was ODD. Therefore, participants in the *Selfish Honesty* condition had the opportunity to either 1) do nothing (engage in omission/inaction) and receive the default allocation, 2) honestly report that the number was odd and receive Option A (*selfish*), or 3) dishonestly report that the number was even and receive Option B (*altruistic*).

All participants made a choice among inaction, Option A, and Option B, and then answered a series of questions about their choice.

Dependent variables.

Choice. Our primary dependent variable was participants' choice of inaction (omission), Option A, or Option B.

Attitudinal measures. Participants also rated how ethical their choice was, the degree to which their choice was motivated by honesty and selfishness, and their moral identity using the same measures we used in Study 1 of the main manuscript ($r_s > .73$, $\alpha_s > .52$).

Emotions. As in Study 1 (see SOM 1.1), participants also rated the extent to which they felt Negative Affect ($\alpha = .84$) and Positive Affect ($\alpha = .89$).

Choice Difficulty. In Study S1, participants also indicated how difficult their choice was; we used a 7-point Likert scale anchored at 1 = "very easy" and 7 = "very difficult".

Preferred Outcome. At the end of the study, we asked participants to choose which of the outcomes (the default allocation, Option A, or Option B) was most attractive to them.

After participants submitted their responses, we collected demographic information and asked participants what they thought the purpose of the study was. Participants then received a bonus payment based upon their decisions.

Results

Preferred outcome. The decision context did not affect the perceived attractiveness of each allocation ($ps > .59$). Table S4, Panel B, and Figure S3, Panel B depicts these results.

Choice. However, the decision context significantly influenced participants' allocation choices. Consistent with H1, we found that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 19.91$, $p < .001$). The coefficients of the multinomial logistic regression reveal that relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation, as revealed by the coefficients of the multinomial logistic regression ($b = 1.61$, 95% CI = [0.90, 2.32]). Table S4, Panel A and Figure S3, Panel A depicts these results.

Table S4. Multinomial logistic regression on choices in Study S1

Panel A: Actual Choice

					95% CI	
Choice	Condition	<i>b</i>	SE	<i>p</i>	Lower	Upper
Altruistic allocation	Selfish Honesty	-14.831	1074.927	0.989	-2121.650	2091.988
	Constant	-1.768	0.409	< .001	-2.569	-0.966
Selfish allocation	Selfish Honesty	1.607	0.360	< .001	0.901	2.312
	Constant	0.137	0.214	0.523	-0.282	0.555

Post-hoc tests

Overall effect of Selfish Honesty $\chi^2 = 19.91, p < .001$

Panel B: Attractive Choice

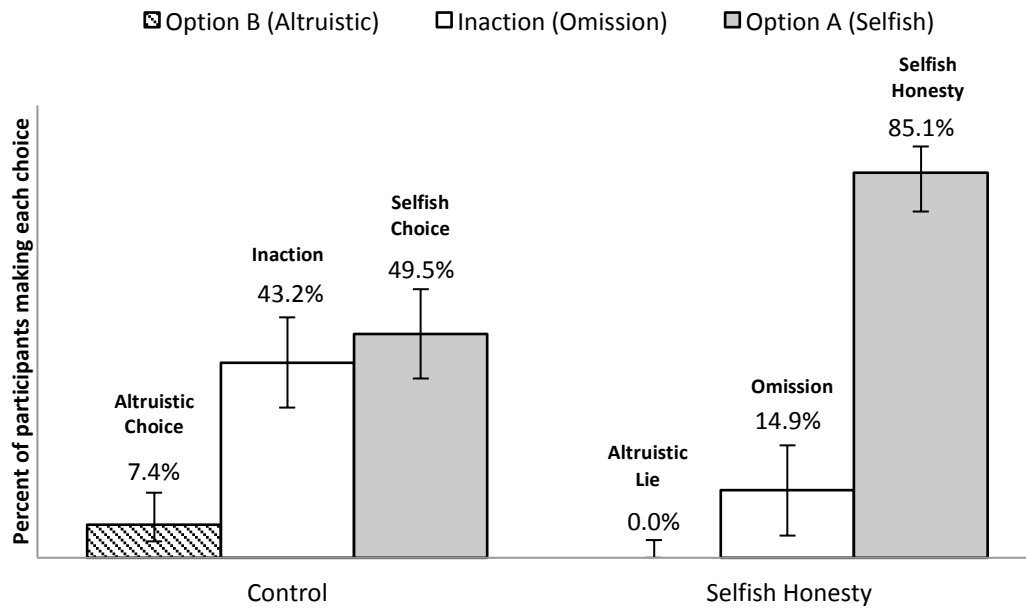
					95% CI	
Choice	Condition	<i>b</i>	SE	<i>p</i>	Lower	Upper
Altruistic allocation	Selfish Honesty	-13.849	759.182	0.985	-1501.819	1474.121
	Constant	-2.673	0.731	< .001	-4.106	-1.241
Selfish allocation	Selfish Honesty	0.170	0.321	0.597	-0.460	0.800
	Constant	0.792	0.224	< .001	0.353	1.230

Post-hoc tests

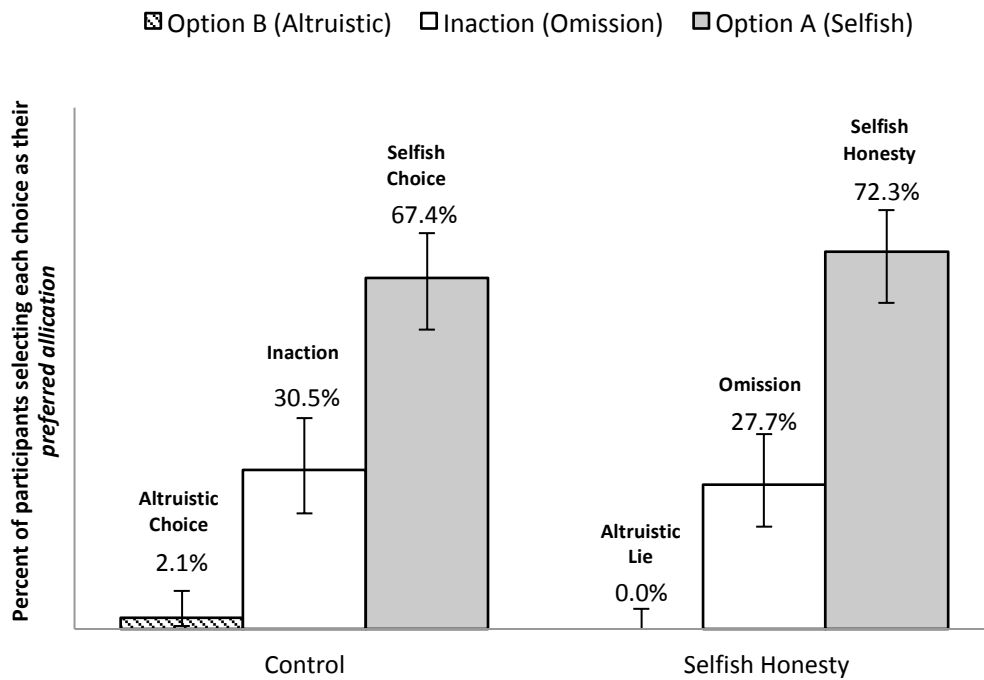
Overall effect of Selfish Honesty $\chi^2 = 0.28, p = .870$

Figure S3. The effect of honesty on selfishness (Study S1)

Panel A: Actual choice



Panel B: Preferred Choice



Note. Error bars reflect 95% confidence intervals.

Attitudinal measures. We intended to conduct 2 (Condition: *Selfish Honesty* vs. *Control*) x 3 (Choice: *Inaction*, *Selfish Allocation*, *Altruistic Allocation*) between-subjects ANOVAs on each attitudinal measure, as indicated in our preregistration. However, because no participants chose the *Altruistic Allocation* in the *Selfish Honesty* condition, these analyses were not valid. Thus, we screened out *Altruistic Allocation* choices (in the *Control* condition) from our main analyses and conducted 2 (Condition: *Selfish Honesty* vs. *Control*) x 2 (Choice: *Inaction*, *Selfish Allocation*) ANOVAs. We present the descriptive statistics and planned contrasts (with corresponding effect sizes) associated with all measures in Table S5.

Ethical Decision. There was a main effect of Condition on the judged ethicality of the decision, ($F(1, 178) = 4.06, p = .045, \eta_p^2 = .022$), such that individuals in the *Selfish Honesty* condition believed their choices were more ethical. There was also a main effect of Choice, $F(1, 178) = 15.21, p < .001, \eta_p^2 = .079$, such that individuals who had chosen the inaction option believed their choices were more ethical.

Importantly, these effects were qualified by a significant Condition x Choice interaction, ($F(1, 178) = 20.33, p < .001, \eta_p^2 = .103$). Individuals who made the selfish choice believed their decision was more ethical in the *Selfish Honesty* condition than in the *Control* condition. However, individuals who chose inaction believed their decision was equally ethical in the *Selfish Honesty* and *Control* conditions.

Justifications.

Honest Justification. There were no main effects of Condition, ($F(1, 178) = 1.95, p = .164, \eta_p^2 = .011$), or Choice, ($F(1, 178) = 1.45, p = .230, \eta_p^2 = .008$), on honest justifications. Importantly, there was a significant Condition x Choice interaction, ($F(1, 178) = 57.19, p < .001, \eta_p^2 = .243$). Individuals who made the selfish choice believed their decision was more motivated by honesty in the *Selfish Honesty* condition than in the *Control* condition. However, individuals who chose inaction believed their decision was less motivated by honesty in the *Selfish Honesty* condition than in the *Control* condition.

Selfish Justification. There was a main effect of Condition on selfish justifications, ($F(1, 178) = 6.49, p = .012, \eta_p^2 = .035$), such that individuals in the *Selfish Honesty* condition believed their choices were more motivated by selfishness than individuals in the *Control* condition. There was also a main effect of Choice, ($F(1, 178) = 99.90, p < .001, \eta_p^2 = .359$), such that individuals who had chosen the *Selfish Allocation* believed their choices were more motivated by selfishness. There was not a significant Condition x Choice interaction, ($F(1, 178) = .011, p = .915, \eta_p^2 < .001$).

Moral Identity. There were no main effects of Condition, ($F(1, 178) = .042, p = .838, \eta_p^2 < .001$), or Choice, ($F(1, 178) = 1.92, p = .168, \eta_p^2 = .011$), on moral identity. Importantly, there was a significant Condition x Choice interaction, ($F(1, 178) = 10.46, p = .001, \eta_p^2 = .056$). Individuals who made the selfish choice saw themselves as higher in moral identity in the *Selfish Honesty* condition than in the *Control* condition. However, individuals who chose inaction saw

themselves as *lower* in moral identity in the *Selfish Honesty* condition than in the *Control* condition.

Emotions.

Positive Affect. There was no significant effect of Condition, ($F(1, 178) = 1.211, p = .273, \eta_p^2 = .007$), or Choice, ($F(1, 178) = 2.09, p = .149, \eta_p^2 = .012$), on positive affect. There was, however, a significant Condition x Choice interaction, ($F(1, 178) = 6.42, p = .012, \eta_p^2 = .035$), such that individuals in the *Control* condition felt more positive affect than individuals in the *Selfish Honesty* condition after engaging in inaction, but did not feel any differently after making the selfish choice.

Negative Affect. There was a marginal effect of Condition, ($F(1, 178) = 3.627, p = .058, \eta_p^2 = .020$), such that individuals felt greater negative affect in the *Control* condition than the *Selfish Honesty* condition. There was also a significant effect of Choice, ($F(1, 178) = 5.280, p = .023, \eta_p^2 = .029$), such that individuals felt more negative emotion after making the selfish choice. There was no significant Condition x Choice interaction, ($F(1, 178) = .121, p = .728, \eta_p^2 = .001$).

Choice Difficulty. There was no significant effect of Condition, ($F(1, 178) = .289, p = .591, \eta_p^2 = .002$), or Choice, ($F(1, 178) = .989, p = .321, \eta_p^2 = .006$), on choice difficulty, nor was there a significant Condition x Choice interaction, ($F(1, 178) = 3.17, p = .077, \eta_p^2 = .017$).

Discussion

Study S1 replicates the effects of Study 1: Individuals are more likely to behave selfishly and more easily justify their behavior when selfishness corresponds with honesty than when it simply reflects a choice.

Table S5. Descriptive statistics of attitudinal measures in Study S1

Condition					Planned contrasts between Control and Selfish Honesty conditions					
Panel A. Ethical decision										
Participant Choice			Control	Selfish Honesty	Total	$ t $	p	d	$Lower\ CI(d)$	$Upper\ CI(d)$
	Inaction(Omission)	M	6.12	5.46	5.96	1.68	0.10	0.52	-0.10	1.13
		SD	1.23	1.37	1.29					
		n	41	14	55					
	Option A (Selfish)	M	3.90	5.63	4.99	6.07	0.00	1.12	0.73	1.50
		SD	1.40	1.62	1.75					
		n	47	80	127					
	Total	M	4.94	5.60	5.28					
		SD	1.73	1.58	1.68					
		n	88	94	182					

Condition						Planned contrasts between Control and Selfish Honesty conditions				
Panel B. Honest justification										
Participant Choice			Control	Selfish Honesty	Total	<i>t</i>	<i>p</i>	<i>d</i>	<i>Lower CI (d)</i>	<i>Upper CI (d)</i>
	Inaction(Omission)	<i>M</i>	5.82	3.57	5.25	5.91	0.00	1.83	1.12	2.52
		<i>SD</i>	1.12	1.52	1.57					
		<i>n</i>	41	14	55					
	Option A (Selfish)	<i>M</i>	4.22	5.77	5.2	5.77	0.00	1.06	0.68	1.44
		<i>SD</i>	1.57	1.38	1.63					
		<i>n</i>	47	80	127					
	Total	<i>M</i>	4.97	5.44	5.21					
		<i>SD</i>	1.59	1.6	1.61					
		<i>n</i>	88	94	182					

Panel C. Selfish justification											
Participant Choice			Control	Selfish Honesty	Total	t/	p	d	Lower CI (d)	Upper CI (d)	
	Inaction(Omission)	M	3.91	3.41	3.78	2.06	0.04	0.63	0.02	1.24	
		SD	0.87	1.16	0.97						
		n	41	14	55						
	Option A (Selfish)	M	5.79	5.33	5.50	1.80	0.07	0.33	-0.03	0.69	
		SD	1.09	1.11	1.12						
		n	47	80	127						
	Total	M	4.92	5.04	4.98						
		SD	1.37	1.31	1.33						
		n	88	94	182						

Panel D. Moral Identity											
Participant Choice	Inaction(Omission)		Control	Selfish Honesty	Total	t/	<i>p</i>	<i>d</i>	<i>Lower CI (d)</i>	<i>Upper CI (d)</i>	
		<i>M</i>	6.24	5.7	6.1	2.80	0.01	0.87	0.23	1.49	
		<i>SD</i>	0.58	0.75	0.66						
	<i>n</i>	41	14	55							
	Option A (Selfish)	<i>M</i>	5.51	5.99	5.81	2.70	0.01	0.50	0.13	0.86	
		<i>SD</i>	1.12	0.86	0.99						
		<i>n</i>	47	80	127						
	Total	<i>M</i>	5.85	5.94	5.9						
		<i>SD</i>	0.98	0.84	0.91						
		<i>n</i>	88	94	182						

...continued

Panel E. Negative Affect											
Choices			Control	Selfish Honesty	Total	t/	p	d	Lower CI (d)	Upper CI (d)	
	Inaction(Omission)	M	2.756	2.5	2.691	0.87	0.39	0.27	-0.34	0.88	
		SD	0.969	0.913	0.953						
		n	41	14	55						
	Option A (Selfish)	M	3.191	2.821	2.958	2.25	0.03	0.41	0.05	0.78	
		SD	0.892	0.898	0.91						
		n	47	80	127						
	Total	M	2.989	2.773	2.877						
		SD	0.949	0.903	0.929						
		n	88	94	182						

Panel F. Positive Affect											
Choices			Control	Selfish Honesty	Total	<i>t</i>	<i>p</i>	<i>d</i>	<i>Lower CI (d)</i>	<i>Upper CI (d)</i>	
	Inaction(Omission)	M	4.274	3.125	3.982	2.18	0.03	0.67	0.05	1.29	
		SD	1.794	1.407	1.765						
		n	41	14	55						
	Option A (Selfish)	M	3.931	4.384	4.217	1.39	0.17	0.26	-0.11	0.62	
		SD	1.601	1.872	1.784						
		n	47	80	127						
	Total	M	4.091	4.197	4.146						
		SD	1.693	1.86	1.777						
		n	88	94	182						

Panel G. Self-reported Choice Difficulty											
Choices		Control	Selfish Honesty	Total	$ t $	p	d	Lower CI (d)	Upper CI (d)		
	Inaction(Omission)	M	2.805	3.214	2.909	0.67	0.5	0.21	-0.4	0.81	
		SD	2.028	1.762	1.956						
		n	41	14	55						
	Option A (Selfish)	M	3.064	2.3	2.583	2.35	0.02	0.43	0.07	0.8	
		SD	1.858	1.717	1.801						
		n	47	80	127						
	Total	M	2.943	2.436	2.681						
		SD	1.932	1.745	1.85						
		n	88	94	182						

3.2. Study S2: Independently manipulating honesty and self-interest

Study S2, like Study 5 of the main manuscript, examines how honesty itself influences behavior when it is associated with neither self-interest nor altruism. However, in Study S2, we independently manipulated honesty and self-interest to tease apart their unique effects in motivating behavior. We use the same advice-giving paradigm we used in Study 5.

Method

Study S2 was preregistered on AsPredicted.org (<https://aspredicted.org/54z5r.pdf>).

Participants. We aimed to recruit 800 adults to participate in an online study via Amazon Mechanical Turk (MTurk), based on the a priori goal of recruiting 200 participants per condition. We ended up with a final sample of 807 adults who passed the comprehension checks (49% female; $M_{\text{age}} = 36$ years, $SD = 15.05$).

Procedure and Materials. In this study, we sought to test how moral and immoral motives (specifically, honesty and selfishness) influence individuals' propensity to look at the social consequences of their actions before speaking up. To do so, we randomly assigned participants to one of four conditions in a 2 (*Honest Motive: Honest, Not Honest*) x 2 (*Selfish Motive: Selfish, Not Selfish*) between-subjects design.

Participants completed an "Advice task," which was nearly identical to the task used in Study 5. However, we varied whether participants honestly knew the amount of money in the jar of coins (*Honest Motive* manipulation) and whether they were personally incentivized to advise their partner that the jar contained a particular amount of money (*Selfish Motive* manipulation).

Participants in the *Not Honest* conditions saw 5 different jars of coins with varying amounts (\$3, \$18, \$23, \$29, and \$50) and were told that their partner would see one of those jars selected at random. Therefore, participants could not use honesty to motivate their behavior because they did not have any insight into what truthful advice would look like. Alternatively, participants in the *Honest* conditions saw the jar of coins that their partner would see and were told that the jar contained exactly \$3. Therefore, participants who were motivated by honesty could simply advise their partner that there were \$3 in the jar.

We also manipulated whether participants could be motivated by selfishness. In the *Not Selfish* condition, participants were told that their bonus payment would not depend on their partner's guess; instead they would get a fixed \$0.10 bonus (as in Study 5). Therefore, participants could not use selfishness to motivate their behavior because there was no personal incentive for any particular action.

Alternatively, in the *Selfish* condition, we informed participants that they would get a \$1 bonus if their partner guessed that there were \$3 in the jar of coins, regardless of whether that was true (as it was in the *Honest* condition) or not (likely not true in the *Not Honest* condition).

Importantly, all participants were told that their partner's bonus payment would depend on their partner's own guess about the amount of money in the jar that they would see. However, they did not know exactly what determined their partner's payment. Participants simply knew that their partner "could receive a bonus payment or a penalty for guessing the correct amount of money in the jar that they see, or guessing a specific amount."

Dependent variables. After reading the information about the task and passing a comprehension check, participants decided which message to send to their partner or whether they would like to learn what determines their partner's bonus payment or not before choosing a message. Participants had the option to send one of ten messages, each indicating that the jar contained a specific dollar amount with values ranging from \$1 to \$10, or the option to "find out what determines my partner's bonus payment before deciding what message to send". Participants who opted to find out their partner's payment learned that their partner would "receive \$1 if s/he guesses that there are \$3 in the jar of coins" and then chose one of the ten messages. The choice to learn what determines their partner's payment was our primary dependent variable, as indicated in our preregistration.

After choosing a message, participants rated their agreement with a series of statements about the importance of truth and how they justified their choice. For these measures we used 7-point Likert scale anchored at 1="Strongly disagree" and 7="Strongly agree." We measured participants' agreement with the importance of truth using three items ($\alpha = .86$): "I care deeply about being honest," "Speaking the truth is the most important value," and "I always tell the truth." We also measured participants' justification for their choice of what message to send. Participants rated their agreement with four distinct justification items: "I chose the easiest course of action" (easy choice justification), "I made my decision because I wanted to avoid an ethical dilemma" (avoid dilemma justification), "I made my decision because it maximized my own payout" (selfish justification), and "I made my choice because it was honest" (honest justification).

Last, we collected demographic information. Participants received their base payment at the conclusion of the study and, within two weeks, received their appropriate bonus payment.

Results

Choice to look. We conducted a logistic regression on the choice to look at the social consequences of one's actions (1 = chose to look, 0 = sent a message without finding out the social consequences), using *Honest Motive*, *Selfish Motive*, and their interaction as independent variables. Figure S4 depicts these results (see "Looked at social consequences" choice).

We found a main effect of *Honest Motive* ($b = -1.272, p < .001$), such that participants who saw the exact jar of coins that their partner would see (and thus, whose behavior could be motivated by honesty) were less likely to seek out information about the social consequences of their action than participants who did not know exactly which jar their partner would see (and thus, whose behavior could not be motivated by honesty). Specifically, only 40.6% of participants with the *Honest* motivation chose to look compared to 65.8% of *Not Honest* participants; $\chi^2(1, N = 805) = 49.93, p < .001$.

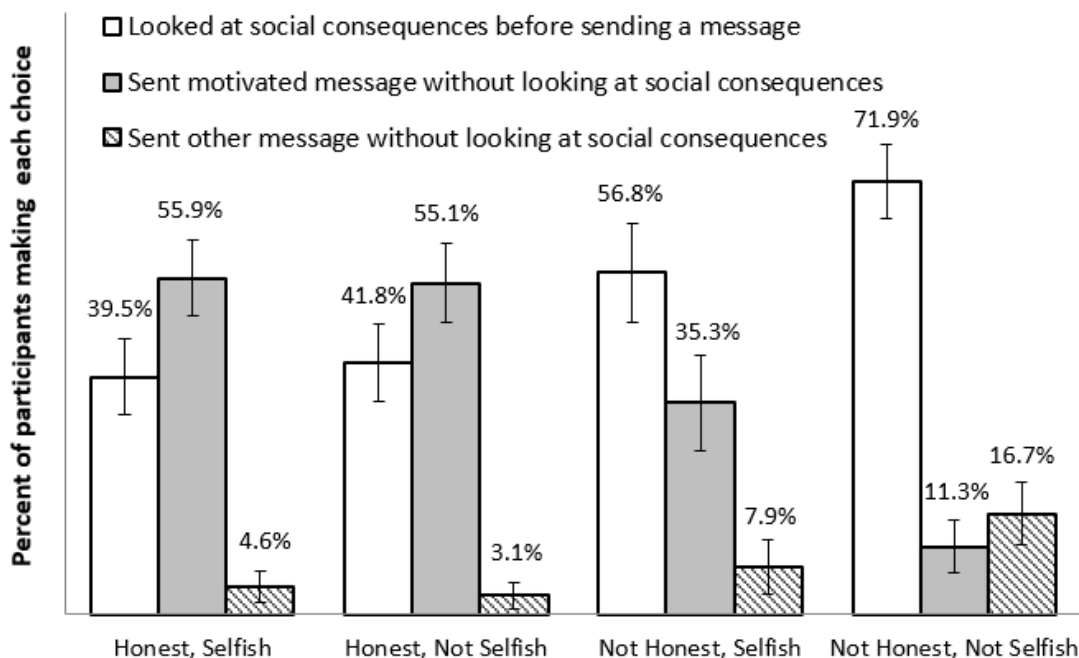
Similarly, we found a main effect of *Selfish Motive* ($b = -0.665$, $p = .004$), such that participants who were incentivized to send a particular message were less likely to seek out information about the social consequences of their action than participants who faced no such incentive. Specifically, only 45.9% of participants with a *Selfish* motivation chose to look compared to 56.1% of *Not Selfish* participants; $\chi^2(1, 805) = 8.32$, $p = .004$.

Interestingly, we found a marginally significant interaction of *Honest Motive* x *Selfish Motive* ($b = 0.571$, $p = .056$). Based on the pattern of means, it appears that the effect of each motivation is diminished in the presence of the other: selfishness influenced the choice to look in the absence of an honest motive (71.9% vs. 56.8%; $\chi^2 = 8.34$, $p = .004$), but not in the presence of an honest motive (41.8% vs 39.5%; $\chi^2 = 0.25$, $p = .61$), and honesty influenced choice more in the absence of a selfish motive (71.9% vs. 41.8%; $\chi^2 = 39.37$, $p < .001$) than in the presence of a selfish motive (56.8% vs 39.5%; $\chi^2 = 10.62$, $p < .001$).

Although we did not have a priori predictions for this interaction, and thus are tentative in over-interpreting this marginal result, this pattern of results suggests the interesting possibility that once individuals have one justification for sending a particular message (or making a particular statement), individuals feel free to act on this justification. They do not need to seek out additional information that would further justify or complicate the decision (e.g., the social consequences of their message) and adding additional justifications does not change behavior.

We found a similar pattern of results when we look at the decision to send the motivated message, rather than the choice to look at the social consequences of one's actions, as the dependent variable.

Figure S4. Results from Study S2



Note: Error bars reflect 95% confidence intervals. Sending the motivated message means sending “There are \$3 in the jar of coins” because this was the message that fulfilled either honest or selfish motives. Sending the other message means sending a message about any other amount from \$1 to \$10.”

Importance of truth. A two-way ANOVA revealed main effects of Honest Motive, $F(1,800) = 19.45, p < .001, \eta_p^2 = .024$, and Selfish Motive, $F(1,800) = 5.47, p = .02, \eta_p^2 = .007$, on the importance of truth. Specifically, *Honest* participants reported valuing truth more than *Not Honest* participants and *Selfish* participants reported valuing truth less than *Not Selfish* participants. Importantly, these effects were qualified by a significant *Honest Motive* x *Selfish Motive* interaction, $F(1,800) = 9.82, p = .002, \eta_p^2 = .012$, such that *Not Honest, Selfish* participants reported valuing truth less than *Not Honest, Not Selfish* participants, whereas *Honest, Selfish* and *Honest, Not Selfish* reported valuing truth similarly. In other words, when individuals had an honest justification available, participants were more likely to believe that honesty was important, regardless of whether they were also motivated by self-interest. However, when individuals only had a selfish justification available, they were less likely to believe that honesty was important, presumably because they recognized that self-interest conflicted with honesty. We report the descriptive statistics in Table S6, Panel A.

Justifications. We also conducted two-way ANOVAs on each self-reported justification measure, using *Honest Motive*, *Selfish Motive* and their interaction as factors. Judgments of the Honest justification followed the same patterns as our importance of truth measure (i.e., main effects of both *Honest Motive* and *Selfish Motive*, as well as an interaction, with the same pattern of means).

For the Avoid Dilemma (“wanted to avoid an ethical dilemma”) justification, we found a main effect of *Honest Motive*, such that *Honest* participants “wanted to avoid an ethical dilemma” more than *Not Honest* participants. Consistent with our theorizing, this result suggests that individuals who have the opportunity to be honest intentionally avoid finding out whether honesty conflicts with other moral principles (e.g., interpersonal harm) in order to circumvent ethical dilemmas.

Confirming the intent of our manipulation, we found a main effect of *Selfish Motive* on selfish justification, such that *Selfish* participants were more motivated by self-interest than *Not Selfish* participants. We also found a main effect of *Selfish Motive* on the Easiest Choice justification, such that *Selfish* participants were more likely to justify their choice as reflecting the “easiest course of action” than *Not Selfish* participants. We present these results in greater detail in Table S6, Panels B-E.

Table S6. Statistics for Importance of Truth and Justification measures in Study S2

Panel A. Importance of Truth					
Selfish Motive	Selfish		Honest Motive		Total
			Honest	Not Honest	
	<i>M</i>	5.68	5.05	5.45	
	<i>SD</i>	1.11	1.35	1.24	
	<i>n</i>	238	139	377	
	Not Selfish	<i>M</i>	5.62	5.51	5.57
		<i>SD</i>	1.19	1.09	1.14
		<i>n</i>	225	202	427
	Total	<i>M</i>	5.65	5.32	5.51
		<i>SD</i>	1.15	1.22	1.19
<i>n</i>		463	341	804	
Main Effect of Honesty Motive		$F(1, 800) = 19.45, p < .001, \eta p^2 = .024$			
Main Effect of Selfish Motive		$F(1, 800) = 5.47, p = .02, \eta p^2 = .007$			
Honest Motive x Selfish Motive		$F(1, 800) = 9.82, p = .002, \eta p^2 = .012$			
Panel B. Easy Choice Justification					
Selfish Motive	Selfish		Honest Motive		Total
			Honest	Not Honest	
	<i>M</i>	4.6	4.58	4.61	
	<i>SD</i>	1.68	1.65	1.66	
	<i>n</i>	238	139	377	
	Not Selfish	<i>M</i>	4.06	3.96	4.01
		<i>SD</i>	1.85	1.82	1.83
		<i>n</i>	224	202	426
	Total	<i>M</i>	4.35	4.21	4.29
		<i>SD</i>	1.79	1.77	1.78
<i>n</i>		462	341	803	
Main Effect of Honesty Motive		$F(1, 799) = 0.36, p = .548, \eta p^2 < .001$			
Main Effect of Selfish Motive		$F(1, 799) = 21.9, p < .001, \eta p^2 = .027$			
Honest Motive x Selfish Motive		$F(1, 799) = 0.03, p = .866, \eta p^2 < .001$			
Panel C. Avoid Dilemma Justification					
Selfish Motive	Selfish		Honest Motive		Total
			Honest	Not Honest	
	<i>M</i>	4.71	4.02	4.45	
	<i>SD</i>	1.79	1.79	1.82	
	<i>n</i>	238	139	377	
	Not Selfish	<i>M</i>	4.43	4.16	4.30
		<i>SD</i>	1.98	1.91	1.95
		<i>n</i>	224	202	426
	Total	<i>M</i>	4.57	4.10	4.37
		<i>SD</i>	1.89	1.86	1.89
<i>n</i>		462	341	803	
Main Effect of Honesty Motive		$F(1, 799) = 12.58, p < .001, \eta p^2 = .016$			
Main Effect of Selfish Motive		$F(1, 799) = 0.25, p = .615, \eta p^2 < .001$			
Honest Motive x Selfish Motive		$F(1, 799) = 2.3, p = .13, \eta p^2 = .003$			

...continued

Panel D. Selfish Justification					
Selfish Motive	Selfish		Honest Motive		
			Honest	Not Honest	Total
		<i>M</i>	4.91	5.08	4.97
		<i>SD</i>	1.71	1.57	1.66
		<i>n</i>	238	139	377
	Not Selfish	<i>M</i>	2.62	2.52	2.57
		<i>SD</i>	1.80	1.76	1.78
		<i>n</i>	224	202	426
	Total	<i>M</i>	3.80	3.57	3.70
		<i>SD</i>	2.09	2.10	2.10
		<i>n</i>	462	341	803
Main Effect of Honesty Motive		$F(1, 799) = 0.1, p = .747, \eta p^2 < .001$			
Main Effect of Selfish Motive		$F(1, 799) = 379.54, p < .001, \eta p^2 = .322$			
Honest Motive x Selfish Motive		$F(1, 799) = 1.12, p = .291, \eta p^2 = .011$			
Panel E. Honest Justification					
Selfish Motive	Selfish		Honest Motive		
			Honest	Not Honest	Total
		<i>M</i>	5.84	4.48	5.34
		<i>SD</i>	1.24	1.73	1.58
		<i>n</i>	238	139	377
	Not Selfish	<i>M</i>	5.97	5.46	5.73
		<i>SD</i>	1.32	1.54	1.45
		<i>n</i>	224	202	426
	Total	<i>M</i>	5.91	5.06	5.55
		<i>SD</i>	1.28	1.69	1.52
		<i>n</i>	462	341	803
Main Effect of Honesty Motive		$F(1, 799) = 82.23, p < .001, \eta p^2 = .093$			
Main Effect of Selfish Motive		$F(1, 799) = 28.66, p < .001, \eta p^2 = .035$			
Honest Motive x Selfish Motive		$F(1, 799) = 16.88, p < .001, \eta p^2 = .021$			

Discussion

In Study S2, we explore how honesty and selfishness independently influence individuals' likelihood of looking at the social consequences of their actions before acting. Importantly, we found that honesty *itself* causes individuals to avoid learning information about how one's action affects others, consistent with Study 5.

However, the results of this study must be interpreted with caution, given differential attribution across our experimental conditions. That is, participants completed the survey in full at different rates in each condition. Participants who did not complete the study in full either dropped out or were unable to pass the comprehension check. The comprehension check comprised three questions: whether their bonus payment depended on their partner's guess ("Yes" for the Selfish condition and "No" for the Not Selfish condition), whether their partner's payment depended on their partner's guess ("Yes" for all participants), and how much money was in the jar of coins (\$3 for the Honest condition and "I don't know" for the Not Honest condition). Participants who failed the comprehension check had the opportunity to reread the information and retake the comprehension check. If any participant failed the comprehension check twice, they were not allowed to complete the study.

Notably, only 51% of participants in the *Not Honest, Selfish* condition passed the comprehension check and completed the survey. We think that the instructions and comprehension check in the *Not Honest, Selfish* condition may have been harder than for the other conditions. Table S7 below provides the completion rates for all conditions.

Table S7. Completion Rates by Condition in Study S2

Condition	N	Complete
Honest, Selfish	272	240 (88%)
Honest, Not Selfish	273	225 (82%)
Not Honest, Selfish	272	139 (51%)
Not Honest, Not Selfish	270	203 (75%)
Not Assigned	10	0 (0%)
Total	1097	807 (74%)

3.3. Study S3: Manipulating whether omission is possible

Study S3 explores how the nature of one's choice set influences self-serving versus costly honesty. In the main manuscript, our experiments modeled situations in which an actor had the opportunity to do nothing. In Study S3, we explore how this inaction (omission) option causally influences selective honesty, as well as the relationship between honesty and both selfishness and altruism.

We hypothesized that omission options have different effects on selfish versus altruistic honesty, and therefore, that selective honesty (H2) would be moderated by the presence of an omission option in one's choice set.

When honesty is costly to a communicator, we posit that omission represents a compromise between self-interest and morality (Levine, Roberts, & Cohen, 2019). Staying silent allows the communicator to feel that they have not explicitly broken a moral rule (which would be violated by lying; Levine et al., 2018), while also not actively sacrificing their own interests. For example, negotiators who omit negative information about a good tend to believe they are more ethical than negotiators who outright lie, and they also reap greater profits than honest negotiators who reveal negative information (Rogers et al., 2016). In general, when honesty is costly to the communicator, communicators are likely to favor omission over honesty (Anderson, 2003; DeScioli, Christner, & Kurzban, 2011; Levine et al., 2018; Pittarello, Rubaltelli, & Motro, 2016).

When inaction is not possible, however, people are less able to reach compromises between self-interest and morality. When honesty is associated with altruism and omission is not possible, people face a rather stark choice between being moral (being both honest and altruistic) and being immoral (being both dishonest and selfish). Most people are honest when they face this choice set (e.g., Capraro, Schulz, & Rand, 2019). Taken together, this leads us to the following hypothesis:

Supplemental Hypothesis 1a (SH1a). People are more likely to engage in altruistic honesty when omission is not possible than when omission is possible.

This hypothesis is consistent with work suggesting that direct questions, which eliminate the possibility of omission (Schweitzer & Croson, 1999), curb selfish deception and promote prosocial honesty (see also Minson, VanEpps, Yip, & Schweitzer, 2018).

Now, we consider how omission options influence communicators' propensity to engage in selfish honesty. When honesty serves a communicators' self-interest, they are not necessarily motivated to identify a moral compromise. Selfish honesty allows a communicator to feel moral *and* to promote their own interests, regardless of whether omission is possible. Therefore, we hypothesize:

SH1b. People are equally likely to engage in selfish honesty, regardless of whether omission is possible.

Taken together, SH1a and SH1b suggest that when omission is possible, people are likely to be honest when honesty is selfish, but to engage in omission when honesty is altruistic. However, when omission is not possible, people may be more likely to be honest, regardless of whether it is selfish or altruistic. This leads us to:

SH1c. Omission options enable selective honesty: People are more likely to exhibit selective honesty when inaction is possible than when inaction is not possible.

Method

Study S3 was preregistered at AsPredicted.org (<https://aspredicted.org/g9kz9.pdf>). However, the analyses (on choice) reported below differ from what we preregistered. We use multinomial (and fifth) logistic regression to analyze choice data, in order to be consistent with the analyses used in the main paper, and to test our hypotheses in the most precise way possible.

Participants. We aimed to recruit 800 adults to participate in an online study via Amazon Mechanical Turk (MTurk). 973 participants started the study, but 161 were automatically kicked out of the study (prior to the manipulation) for failing an attention check, and an additional 16 left the survey before completing our main measures. We ended up with a final sample of 796 adults (789 of whom provided demographic information; 46.5% female; $M_{\text{age}} = 38$ years, $SD = 12.36$) who were eligible for analysis. For all analyses, we include all participants who responded to the relevant dependent measure.

Procedure and materials. In Study S3, we randomly assigned participants to a condition from a 3 (Decision context: *Control*, *Selfish Honesty*, *Altruistic Honesty*) x 2 (Omission Possible: *Omission*, *No Omission*) between-subjects design. The *Omission* conditions were identical to the conditions in Study 1; participants were in the role of Decider and had to decide whether to do nothing and keep the current allocation as is, or switch it to a more selfish (Option A) or altruistic (Option B) allocation, and we varied whether Option A or Option B was associated with Honesty.

In the *No Omission* conditions, we simply removed the possibility of inaction (omission). Therefore, within these conditions, participants in the *Control* condition had to choose between Option A (selfish) and Option B (altruistic), participants in the *Selfish Honesty* condition had to choose between telling the truth which resulted in Option A (selfish) or lying which resulted in Option B (altruistic), and participants in the *Altruistic Honesty* condition had to choose between lying which resulted in Option A (selfish) or telling the truth which resulted in Option B (altruistic). Participants were not presented with an initial allocation, and could not choose inaction.

Dependent variables. The primary dependent variable was participants' distribution of choices across conditions. Participants also rated how ethical their choice was using the same *Ethical Decision* scale used in Study 1 ($r(795) = .93$, $p < .001$).

After participants submitted their responses, we collected demographic information, and asked participants what they thought the purpose of the study was. Participants then received a bonus payment based upon their decisions.

Results

To test SH1a-c, we conducted a first logistic regression on honest choice, using Omission Possible condition (1 = *Omission*, 0 = *No Omission*), Decision Context condition (1 = *Selfish Honesty*, 0 = *Altruistic Honesty*), and their interaction as independent variables. This analysis necessarily omitted the *Control* condition.

We then ran supplementary to explore what underlies (the potential) differential rates of selective honesty when omission is or is not possible. Specifically, within the *Omission* condition, and within the *No Omission* condition, we tested whether selective honesty was best explained by genuine or motivated preferences for honesty. We used multinomial logistic regression (identical to the analyses used in Study 1) within the *Omission* condition, and we used logistic regression (identical to the analyses used in Study 4) within the *No Omission* condition.

We depict the full pattern of results across conditions in Figure S5.

First logistic regression comparing selective honesty in the *Omission* and *No Omission* conditions. We found a main effect of Decision Context ($b = 2.65$, 95% CI = [1.52, 3.77]), such that participants were more likely to be honest in the *Selfish Honesty* condition than in the *Altruistic Honesty* condition. We also found a main effect of Omission Possible ($b = -2.62$, 95% CI = [-3.22, -2.01]) such that participants were less likely to be honest when omission was possible than when omission was not possible. Although the effect of *Selfish Honesty* was qualitatively larger in the *Omission possible* condition ($\chi^2 = 59.79$, $p < .001$) than in the *Omission not possible* condition ($\chi^2 = 33.08$, $p < .001$), we did not have sufficient power to detect a Decision Context x Omission Possible interaction ($b = -2.51$, 95% CI = [-1.78, .75]). In other words, we did not find evidence for SH1c.

Choice within the *Omission* condition.

We found that participants' distribution of choices was significantly different in the *Selfish Honesty* condition and the *Control* condition ($\chi^2 = 27.54$, $p < .001$). This difference occurred primarily because relative to the *Control* condition, participants in the *Selfish Honesty* condition shifted their choices away from inaction and towards the selfish allocation, as revealed by the coefficients of the multinomial logistic regression ($b = 1.37$, 95% CI = [0.88, 1.89]).

We also found that participants' distribution of choices was significantly different in the *Altruistic Honesty* condition and the *Control* condition ($\chi^2 = 11.08$, $p = .004$). Relative to the *Control* condition, participants in the *Altruistic Honesty* condition shifted their choices away from inaction and towards the altruistic allocation, ($b = 2.11$, 95% CI = [0.87, 3.36]).

Though having the opportunity to tell a selfish truth (*Selfish Honesty* condition) had a qualitatively larger effect on participants' distribution of choice than having the opportunity to

tell an altruistic truth (*Altruistic Honesty* condition) did, the effect of *Selfish Honesty* on selfish choice was not significantly different than the effect of *Altruistic Honesty* on altruistic choice ($\chi^2 = 1.21, p = .272$). In other words, we are not able to reject the possibility that honesty enables selfishness and selflessness to a similar degree.

Choice within the *No Omission* condition.

We found a main effect of *Selfish Honesty* ($b = 2.06, 95\% \text{ CI} = [0.91, 3.21]$), such that participants were more likely to be selfish when selfishness was associated with honesty (i.e., *Selfish Honest* vs. *Control*). We also found a main effect of *Altruistic Honesty* ($b = -2.51, 95\% \text{ CI} = [-3.10, -1.92]$), such that participants were less likely to be selfish (and more likely to be altruistic) when altruism was associated with honesty (i.e., *Altruistic Honest* vs. *Control*). These effect sizes (in absolute value) are not significantly different ($\chi^2 = .37, p = .54$). In other words, when omission is not possible, honesty enables selfishness and selflessness to a similar degree (see Table S8, Panel B for further details).

Table S8. Logistic regressions on choices in Study S3

Panel A: Omission is Possible (Multinomial Logistic Regression)

					95% CI	
Choice	Condition	<i>b</i>	SE	<i>p</i>	Lower	Upper
Altruistic allocation	Altruistic Honesty	2.113	0.636	0.001	0.866	3.360
	Selfish Honesty	0.214	0.930	0.818	-1.609	2.036
	Constant	-3.412	0.587	< .001	-4.562	-2.262
Selfish allocation	Altruistic Honesty	0.059	0.280	0.833	-0.489	0.607
	Selfish Honesty	1.374	0.263	< .001	0.859	1.889
	Constant	-0.847	0.191	< .001	-1.222	-0.472

Post-hoc tests

Overall effect of Altruistic Honesty

$\chi^2 = 11.08, p = .004$

Overall effect of Selfish Honesty

$\chi^2 = 27.54, p < .001$

Effect of Altruistic Honesty on Altruism vs. Effect of Selfish Honesty on Selfishness

$\chi^2 = 1.21, p = .272$

Panel B: Omission is Not Possible (Firth Logistic Regression)

					95% CI	
Choice	Condition	<i>b</i>	SE	<i>p</i>	Lower	Upper
Selfish allocation	Altruistic Honesty	-2.510	0.299	< .001	-3.096	-1.924
	Selfish Honesty	2.063	0.587	< .001	0.912	3.214
	Constant	1.548	0.227	< .001	1.103	1.993

Post-hoc test

Effect of Altruistic Honesty on Altruism vs. Effect of Selfish Honesty on Selfishness

$\chi^2 = .37, p = .54$