I didn't know either: how beliefs about norms shape strategic ignorance

Hua, Tony

University of California, Merced

1 April 2025

I Didn't Know Either: How Beliefs About Norms Shape Strategic Ignorance*

Tony Hua[1]

[1]Department of Economics and Business Management, School of Social Sciences, Humanities & Arts, University of California, Merced, 5200 N. Lake Rd., Merced, CA 95343, United States of America

July 2025

Abstract

People often avoid information as a way to justify selfish behavior. However, such behavior often unfolds in social contexts, where expectations about others' behavior may shape moral decision-making. This study investigates how beliefs about descriptive norms influence strategic ignorance in a modified moral wiggle-room game. Participants first predicted how often others acquired information, then received randomly assigned feedback indicating high or low rates of ignorance before making their own decision as the dictator. Individuals were more likely to seek information when told that most others typically did so, and more likely to avoid information when told that ignorance was common. Norm-conforming behavior differed between *ignorance expectants*–those who expected ignorance and *reveal expectants*–those who expected information acquisition. *Ignorance expectants* adjusted only when norm cues strongly contradicted their prior beliefs, whereas *reveal expectants* exhibited modest but consistent responsiveness across norm environments. There is no evidence of ex-ante self-serving belief distortion, though limited behavioral change constrained opportunities to test for ex-post justification. These findings suggest that strategic ignorance responds to descriptive norms, but not through motivated belief distortion.

Keywords: information avoidance, moral wiggle-room, belief formation, social norms, social appropriateness, experiment

JEL Codes: C72, C91, D8, D9

*Email: thua5@ucmerced.edu. ORCID: 0009-0002-7815-0386. I am very thankful to Steve Seltzer for allowing me access to his class as subject pools for the pilot studies of this paper and to my advisor, Zack Grossman, for our frequent discussions. I would also like to thank Paul Smaldino, Ignacio Gonzalez Correa, Karine Nyborg, Rob Innes, Ana Tur Prats, and Richard Patterson for their helpful feedback and comments. Funding for the experiments was made possible in part by the Economics Department at the University of California, Merced.

1 Introduction

A well-documented feature of decision-making is that individuals strategically avoid information to justify self-interested behavior. This phenomenon, often referred to as strategic or willful ignorance, is one way individuals create moral wiggle room, enabling them to act selfishly while preserving a positive social or self-image (Dana, Weber, and Kuang, 2007; Feiler, 2014; Z. Grossman, 2014; Z. Grossman and Van der Weele, 2017; Exley and Kessler, 2023). Such behavior might involve ignoring the ethical implications of a purchase, avoiding news about harmful industry practices, or choosing not to learn how one's decisions affect others. While much of the literature emphasizes internalized moral concerns, decision-making rarely occurs in a social vacuum. Individuals frequently assess their choices against perceived norms. For example, someone may reason, "I am not a bad person for not knowing because there are many others who didn't know either." Such reasoning reflects the potential for coordination on norms of either information avoidance or information acquisition, depending on what individuals expect others to do.

There are two main channels for norm compliance. One possibility is that individuals act based on pre-existing beliefs about what others do. Alternatively, people may form beliefs about others' behavior in ways that justify their own actions (Golman, Hagmann, and Loewenstein, 2017). This raises the question of how individuals respond when exposed to information that contradicts their perceptions of others, or when they move between environments with different normative expectations. Deviating from the norm can cause psychological discomfort, while conforming may offer reassurance or utility through social alignment (Akerlof, 1980). These dynamics suggest that strategic ignorance may not only serve a psychological self-justification function, but may also be shaped and reinforced by evolving beliefs about what others consider typical or acceptable. In turn, such belief updating can facilitate coordination, leading individuals to converge, or "pool," on shared behaviors such as avoiding or acquiring information, even in the absence of direct communication or formal incentives.

Extending this line of reasoning, this paper examines how individuals' beliefs about the prevalence of information avoidance—i.e., descriptive norms—influence their own tendency to remain strategically ignorant. I investigate both how prior beliefs about others' behavior shape one's decision to seek or avoid information, and how exposure to updated norm information influences that decision. While social norms are known to shape economic behavior in diverse settings, including tax compliance, market participa-

tion, and prosocial conduct (Tomasello, Kruger, and Ratner, 1993), less is known about whether individuals adopt self-serving beliefs about these norms to rationalize ignorance. For instance, people might excuse their choices by assuming that information avoidance is widespread, or reinterpret what is socially appropriate (i.e., injunctive norms) in a way that justifies avoidance (Bicchieri, Dimant, Gelfand, et al., 2023). Such justifications may contribute to persistent ethical blind spots in economic decision-making. This paper addresses two key questions: (1) Are individuals' information-avoidance decisions influenced by observed social behavior? (2) Do individuals form self-serving beliefs about others' tendency to avoid information?

Using a controlled experiment, I test how participants' prior expectations about others' behavior interact with experimentally provided norm feedback to shape their own information avoidance. Participants engage in a modified moral wiggle-room game (Dana, Weber, and Kuang, 2007), in which a dictator makes a binary allocation decision affecting both themselves and a recipient. The dictator initially lacks information about the recipient's payoff but can reveal it at no cost. Before making their decision, dictators first predict the proportion of other participants from separate experiments who had chosen to avoid information. They are then randomly assigned to one of four norm conditions: the *No Norm* condition (no feedback), the *10% Ignorance* condition (told that only 10% avoided information), the *30% Ignorance* condition (30% avoided), or the *70% Ignorance* condition (70% avoided). This variation in norm feedback allows for a clean test of my first research question: how observed social behavior shapes individual information avoidance.

The experiment also includes a framing manipulation: some participants are told from the outset that they will make the allocation decision themselves (*Known*), while others are initially led to reason about someone else's decision before learning that they will assume that role (*Unknown*).[1] This manipulation allows me to test my second research question: whether individuals form self-serving beliefs, specifically, whether learning they will make the decision after first predicting others' behavior leads them to distort those beliefs in a way that justifies their own potential ignorance. Together, this design enables a clean test of how both self-generated expectations and externally provided norms influence strategic ignorance. Section 2 further describes the experimental procedures.

The results show that both pre-existing beliefs about norms and experimentally pro-

[1] See also Gneezy et al. (2020), who show that self-deception is constrained by the timing of incentive information when advisors evaluate options that affect others.

vided norm feedback influence strategic ignorance. Participants were more likely to seek information when exposed to norms favoring information acquisition, with *ignorance expectants*—those who initially believed avoidance was common—showing increased responsiveness, but only when exposed to a strongly opposing norm (the *10% ignorance* environment). Furthermore, there is no evidence that participants distorted their beliefs in a self-serving manner, either before or after making their decisions. Participants' own views about what was appropriate remained stable across conditions, suggesting that behavior reflected sensitivity to norms but was grounded in personal values.

For social norms to facilitate coordination, individuals must be responsive to others' behavior. While some participants adjusted their choices based on norm feedback, there was little evidence of broad convergence or "pooling" toward either universal ignorance or universal information-seeking. Instead, norm effects were concentrated among specific belief types. *Ignorance expectants* changed behavior when feedback sharply contradicted their expectations, whereas *reveal expectants*—those expecting widespread information acquisition—showed modest and consistent shifts in response to norm cues. These dynamics suggest that convergence toward an ignorance norm is more likely in communities where individuals tend to expect that others will also choose to remain ignorant.

These findings advance our understanding of how social norms shape economic decision-making by identifying conditions under which individuals engage in strategic ignorance. Specifically, the results show that information avoidance is responsive to beliefs about others' behavior, highlighting the potential of norm-based policy interventions to reduce such avoidance. More broadly, the study contributes to ongoing debates on moral decision-making, norm compliance, belief formation, and behavioral interventions in economic contexts. These themes are explored in detail in Section 3. The next section presents the 2×4 experimental design.

2 Experiment Design

This section details the experimental design, key treatment manipulations, and hypotheses. Figure 1 summarizes the experimental procedures. The experiment employs the moral wiggle-room game as its primary decision-making instrument, as it encapsulates the essential features of an information avoidance environment while facilitating comparisons with previous studies (Dana, Weber, and Kuang, 2007; Vu et al., 2023). In this game, participants act as dictators, choosing an allocation of experimental currency

units (ECUs) between themselves and a recipient. The dictator chooses between two possible allocations. The dictator always receives either 6 or 5 ECUs, while the recipient's payoff—randomly determined with equal probability—remains hidden and can be either 1 or 5 ECUs, creating two possible game types.

In the *aligned interest game* (50% probability), the dictator chooses between a 6-5 allocation (where they receive 6 ECUs and the recipient gets 5) or a 5-1 allocation (where they receive 5 ECUs and the recipient gets 1). In the *conflicting interest game* (50% probability), the dictator chooses between a 6-1 allocation (keeping 6 ECUs while the recipient gets 1) or a 5-5 split (where both receive 5 ECUs), as shown in Table 1. Initially, the dictator is uncertain which game type is active but can costlessly reveal this information before making their decision.

Table 1: The two payoffs tables featured in the moral wiggle-room game (the conflicting and aligned states) matched those used by Dana, Weber, and Kuang (2007) and were equally likely.

	Conflicting Interest		**Aligned Interest**	
Dictator chooses	**Dictator gets**	**Recipient gets**	**Dictator gets**	**Recipient gets**
A	6	1	6	5
B	5	5	5	1

2.1 Instructions

In the first section of the experiment, subjects are introduced to the moral wiggle-room game (Dana, Weber, and Kuang, 2007), told that they will be predicting the behavior of what other people are choosing to do in the scenario, and complete comprehension checks. Subjects who failed the checks three times or let the timer expire were removed from the study.[2] In the *Known* condition, subjects are told from the outset that they will be making the allocation decision, encouraging them to form beliefs and expectations from a personal perspective.[3] In the *Unknown* condition, subjects initially reason about what others (referred to as "Person 1") would do before discovering that they will take on that role themselves later in the dictator game stage. This framing variation is designed to test whether early role salience increases the psychological moti-

[2]Depending on the screen, subjects had between 2 to 3 minutes to complete the comprehension check.

[3]In the *Known* condition, the role of "Person 1" was made synonymous with "you." This was intended to reinforce and remind the subject of their pending role as the dictator. Appendix C describes in further details the effects of this *self-referential* framing.

vation to form self-serving beliefs about others' behavior, thereby addressing the second research question on whether belief formation itself is influenced by the anticipation of morally charged decisions.

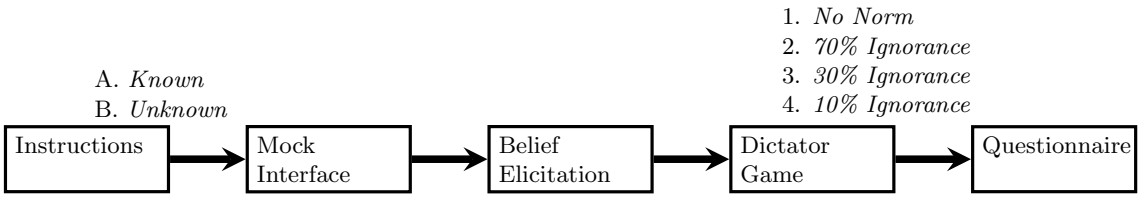


Figure 1: Overview of experimental procedures, with treatment variations indicated above each stage.

2.2 Mock-Up Allocation and Belief Elicitation

After successfully completing the comprehension checks, subjects first preview the allocation interface to become familiar with the decision environment and to begin processing how they might act before moving on to predict the percentage of dictators who revealed payoff tables in a previous study. Specifically, subjects answer: "What percentage of people acting as Person 1 (the dictator) do you believe revealed the payoff tables?"[4] The mock interface familiarizes participants with the decision environment and helps ensure that their beliefs reflect informed predictions. To encourage thoughtful predictions, in the belief elicitation stage, participants were incentivized with a bonus ECU for making a more accurate estimate than the majority of other participants.[5]

The belief question refers to behavior in a mock-up version of the moral wiggle-room game, described as similar to what others previously faced. This deliberate simplification balances methodological clarity and theoretical accuracy. Since descriptive norms are based on what people *believe* others typically do—rather than on the exact details of how those behaviors occurred—leaving out specific procedural elements keeps the task simpler without compromising the purpose of measuring perceived norms. By describing the prior experiments as "similar," the study preserved ambiguity necessary for authentic belief elicitation, while remaining truthful in substance.

[4]The experiment avoided descriptive labels for the role of dictator. Instead, dictators were referred to as "Person 1."

[5]To test whether incentives biased responses, a separate trial elicited beliefs about the same information decision without any bonuses. The belief distributions in the incentivized and non-incentivized conditions are statistically indistinguishable ($D = 0.0548$, $p = 0.901$; Kolmogorov–Smirnov test), suggesting that the bonus did not influence reported beliefs.

2.3 Dictator Decision

After making their predictions, participants are assigned the role of dictators and proceed to the allocation stage of a modified moral wiggle-room game. Before making their allocation decision, they can choose to reveal the recipient's payoff table at no cost.[6] They are informed that recipients will not know whether the dictator chose to reveal the information.

To examine how observed social behavior influences individual information avoidance—a central focus of the first research question—participants receive norm information that varies across four between-subjects conditions. In the *No Norm* condition, no social feedback is provided. In the remaining *10% Ignorance*, *30% Ignorance*, and *70% Ignorance* treatments, participants are shown feedback indicating the rate at which others chose to avoid information in a similar study. This message includes their own prior estimate, whether they overestimated or underestimated the true value, and a norm cue framed as: "A majority (minority): X% or N out of 10 people chose to reveal the payoff tables," where X is the percentage and N is the frequency.

The ignorance rates used in this study are drawn from two prior experiments that demonstrated how subject pool, recipient type, and choice architecture significantly influence information avoidance (Lind, Nyborg, and Pauls, 2019; Z. Grossman, Hua, et al., 2025). The *30% Ignorance* condition is based on Z. Grossman, Hua, et al. (2025), where a shift in choice architecture, requiring subjects to decide on acquiring information before making an allocation decision, substantially reduced ignorance. The *70% Ignorance* condition reflects another treatment from the same study, closely mirroring the canonical design of Dana, Weber, and Kuang (2007), where subjects choose whether to acquire information before making an allocation decision. Lastly, the *10% Ignorance* condition is based on Lind, Nyborg, and Pauls (2019), where differences in the subject pool, recipient type, and subtle differences in the choice architecture resulted in a notably low rate of ignorance.

Finally, the selected norm cues—*30%* and *70% Ignorance*—were chosen to clearly signal directional trends, enabling a strong test of norm-following behavior. The *10% Ignorance* condition was added based on pilot data, where participants typically estimated ignorance rates near 20%, in order to examine how subjects respond when they

[6] Unbeknownst to them, each subject also serves as a recipient for another dictator in the same treatment arm, revealed only after all decisions are made to simplify logistics and reduce costs.

substantially overestimate the norm.[7]

2.4 Questionnaire

Finally, to assess injunctive norms, participants rate the social appropriateness of choosing to reveal or not revealing the recipient's payoff tables. Their responses are incentivized with a 1 ECU bonus if they match the most common answer given by other participants, following the coordination method outlined by Krupka and Weber (2013). After this task, participants complete a questionnaire battery, which includes:

1. The Conformity Scale (Mehrabian and Stefl, 1995): A psychological measure assessing individuals' tendency to conform to social expectations and group behavior. Higher scores indicate a stronger inclination to follow social norms and external influences.

2. A general demographics questionnaire: Collecting information on age, gender, education, and other personal characteristics.

3. An optional Moral Universalism Trust Survey (Enke, Rodríguez-Padilla, and Zimmermann, 2022): This survey evaluates the extent to which individuals place trust in others across different social distances, measuring whether they exhibit in-group favoritism or a more universalist moral perspective—which may influence ethical decision-making and social behavior.

2.5 Hypotheses

Building on Spiekermann and Weiss (2016), who model subjective norm compliance as contingent on beliefs, this study investigates whether perceived descriptive norms influence strategic ignorance via mechanisms of social coordination or self-justification. I test four pre-registered hypotheses that explore how individuals adjust their information avoidance in response to norm cues, prior beliefs, and self-serving motivations. Together, these hypotheses assess the broader claim that multiple, possibly conflicting, ignorance norms can coexist, driven by heterogeneous beliefs and sensitivity to the normative environment. All hypotheses were pre-registered with the AEA RCT Registry.

Social norms play a critical role in shaping behavior. When individuals observe that most others acquire information, they may feel pressure to conform to this norm

[7] See Appendix B for further details.

of information acquisition. Conversely, when ignorance is perceived as common, they may feel more justified in avoiding information themselves. Hypothesis 1 tests this prediction by comparing ignorance rates across conditions: individuals should be more likely to avoid information in the *70% Ignorance* condition than in the *30%* or *10% Ignorance* environments.

Hypothesis 1 - Norm Following: High (low) ignorance norms will lead to higher (lower) propensity to avoid information.

While social norms influence behavior, individuals may respond differently depending on how much their prior beliefs deviate from the observed norm. Those who severely underestimated or overestimated the extent of ignorance may be particularly sensitive to learning about the social norm. To test Hypothesis 2, I examine how the gap between participants' prior beliefs and the normative ignorance rates they are exposed to affects their subsequent choices. In particular, I compare behavior across subjects classified as *reveal expectants* (whose prior expected more than 50% revealed) and *ignorance expectants* (whose prior expected 50% or less to reveal) to determine whether larger deviations from the observed norm produce stronger behavioral adjustments.

Hypothesis 2 - Norm Sensitivity: The further a dictator's priors are below (above) the observed ignorance rate, the more likely they will be to avoid (acquire) information.

People may use ignorance as a moral justification for self-serving behavior. When individuals initially underestimate how common ignorance is, learning that others also avoid information may legitimize their own decision to remain ignorant. In contrast, overestimating ignorance may not create the same pressure to adjust behavior. This asymmetry suggests a stronger incentive to adopt ignorance when it is underestimated than to seek information when it is overestimated. To test this, I compare the behavior of individuals who underestimated the observed ignorance rate to those who overestimated it.

Hypothesis 3 - Exculpatory Norms: When dictators underestimate ignorance norms, they are more likely to avoid information than when they overestimate them.

Beyond norm-driven behavior, individuals may also engage in self-serving belief formation—the tendency to process information in ways that justify their preferred actions. Often occurring subconsciously, this bias allows people to preserve a positive self-image even when acting selfishly (Shepperd, Malone, and Sweeny, 2008). For example, self-serving interpretations distort fairness judgments in negotiations (Babcock and Loewen-

stein, 1997), and motivated reasoning enables individuals to justify self-interested choices without abandoning moral standards (Rustichini and Villeval, 2014; Gino, Norton, and Weber, 2016). Recent work by Bicchieri, Dimant, and Sonderegger (2023) and Exley and Kessler (2024) further shows that people selectively distort or reinterpret morally relevant information to rationalize selfish behavior. These findings suggest that strategic ignorance may not merely reflect norm compliance or moral evasion, but also active distortion of beliefs to align one's choices with a favorable moral narrative.

In other words, if individuals are motivated to behave selfishly without guilt, one might expect them to distort their perception of others' behavior to justify their own information avoidance. Such motivated cognition can lead individuals to inflate perceived ignorance norms, thereby legitimizing their own inaction and reducing the psychological costs of selfish decisions. To test this, I compare participants' beliefs about how often others choose to acquire information between two groups: those who already know they will make the allocation decision (the *Known* condition) and those who have not yet been told they will take on that role (the *Unknown* condition). If self-serving belief formation occurs, those in the *Known* condition should report higher expected ignorance rates.

Hypothesis 4 - Self-serving Beliefs: When dictators anticipate having to exploit moral wiggle-room, they will predict that others are less likely to acquire information.

Hypotheses 1 through 3 address the first research question by investigating whether individuals' information avoidance decisions are influenced by observed social behavior. They focus on the role of descriptive norms in shaping strategic ignorance. In contrast, Hypothesis 4 examines the second research question, asking whether individuals form self-serving beliefs about others' tendency to avoid information to justify their own ignorance. The next section presents the findings.

3 Results

This section begins by outlining the procedures for participant recruitment and data collection. It then describes the data cleaning process, including exclusions and adjustments made to ensure data quality. Next, the analysis verifies whether participants' behavior aligns with theoretical expectations before systematically presenting the key findings. The results are structured to first examine overall patterns, followed by tests of the study's main hypotheses, and concluding with exploratory analyses of unexpected

findings.

3.1 Data

The study was pre-registered under AEARCT-0014367,[8] and the experimental interface was programmed using LIONESS (Giamattei et al., 2020). Participants were recruited via Prolific in January and February of 2025. The sample was restricted to English-speaking subjects in the United States from the age of 19 to 80 with at least 10 completed studies on their Prolific profile with an approval rating of 99-100%. Recruitment was set to retrieve an even split of male and female subjects. To mitigate risks of sample imbalance from varying subject availability, all treatment arms ran simultaneously with subjects being randomly assigned to one of eight treatments in a 2×4 factorial design. A hidden timer (15–16 minutes) regulated study duration, and participants received a base payment of \$2.50–\$2.70 for completing the study.[9] Additional earnings from decision-making tasks were provided through bonus payments at the study's conclusion at a rate of 2 ECUs = \$1.

Two participants who revealed the payoff tables but did not make an allocation decision were excluded from the analysis, resulting in a final sample of 1,020. Additionally, 107 participants did not complete the questionnaire before the timer expired. Partial demographic data were recovered for 62 of them using Prolific user records.[10] The sample was qualitatively balanced across treatment arms, consisting primarily of adults averaging between 37-39 years old, with most identifying as White and only a small proportion currently enrolled as students. Table 2 presents a simplified demographic breakdown, with a full breakdown available in Appendix A.

3.2 Primary Analysis

To confirm that subjects behaved in line with previous studies, I examined dictators' allocation decisions. Among the 370 who remained ignorant, 85% chose the self-serving option (6 over 5), consistent with past evidence that ignorance facilitates selfish behavior. Informed dictators in the aligned-interest game overwhelmingly (98%) chose the

[8] Link to pre-registration: https://doi.org/10.1257/rct.14367-1.2

[9] In the second half of data collection, the timer was extended from 15 to 16 minutes to allow more time for the final questionnaire. The base payment was adjusted accordingly to maintain a consistent advertised rate on Prolific.

[10] For the remaining 45 participants, a technical error–mostly occurring in the first data collection batch–prevented data retrieval.

Table 2: Abridged subject demographics breakdown across treatment groups. Means are reported, with standard deviations in parentheses. Demographic characteristics are balanced across conditions, supporting internal validity of treatment effects.

	Treatment Arms								
	Unk. N/A n=132	Unk. 70% Ig. n=119	Unk. 30% Ig. n=119	Unk. 10% Ig. n=134	Kwn. N/A n=127	Kwn. 70% Ig. n=134	Kwn. 30% Ig. n=135	Kwn. 10% Ig. n=120	Total n=1,020
age									
	38.30	38.27	38.80	39.02	39.11	37.12	37.33	38.13	38.25
	(12.10)	(12.48)	(11.10)	(11.44)	(13.67)	(12.73)	(12.51)	(11.60)	(12.22)
female									
	0.55	0.51	0.58	0.51	0.52	0.42	0.46	0.44	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
student									
	0.17	0.17	0.06	0.10	0.12	0.16	0.12	0.12	0.13
	(0.38)	(0.37)	(0.24)	(0.30)	(0.33)	(0.37)	(0.33)	(0.33)	(0.33)
white									
	0.75	0.73	0.69	0.66	0.71	0.71	0.75	0.77	0.72
	(0.44)	(0.44)	(0.46)	(0.48)	(0.45)	(0.46)	(0.44)	(0.43)	(0.45)

prosocial option (6-5), while in the conflicting-interest game, 78% selected the fairer split (5-5) over the selfish one (6-1). These results replicate standard patterns from moral wiggle-room experiments, where access to information tends to reduce selfishness. With this validation, I now turn to the main hypotheses.

First, I look at how norms about information avoidance affected the dictator's behavior. Figure 2 shows the ignorance rates across the four information treatments. Dictators were 8-12 percentage points more likely to avoid information in the high ignorance (*70%*) environment compared to the low ignorance (*30%* and *10%*) environments respectively. This corresponds to a relative reduction in ignorance behavior of approximately 20–30%, suggesting that even subtle social information cues can meaningfully shift decision-making in morally ambiguous contexts.

A nonparametric Cuzick trend test confirms this pattern ($z = -2.69, p = 0.0071$), indicating a statistically significant monotonic decline in ignorance rates from high to low ignorance norms. Appendix E shows this result is robust to demographic controls. This supports *Hypothesis 1 - Norm Following*, indicating that the decision to avoid information is influenced by the prevailing social norms.

Result 1: Dictators' willingness to avoid information is sensitive to norms about information acquisition.

To test Hypothesis 2 – Norm Sensitivity, I examine whether the distance between participants' prior beliefs and the observed norm feedback influences their decisions. Because prior beliefs about descriptive norms are endogenous, I introduce exogenous

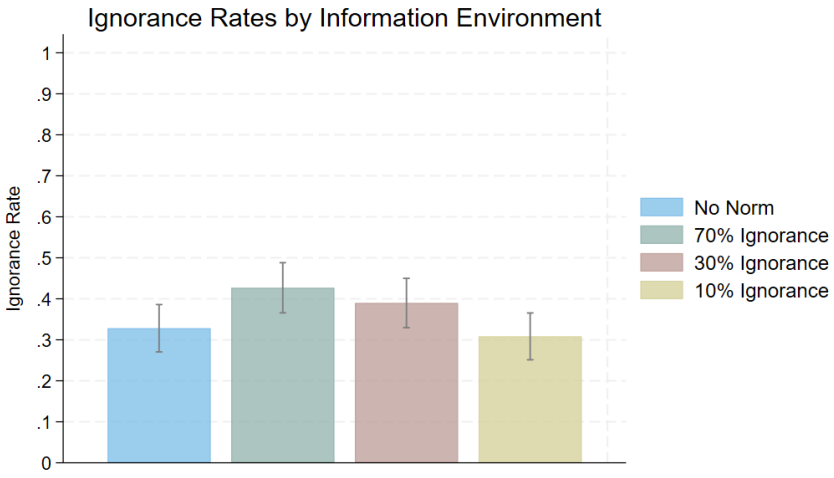


Figure 2: Information avoidance rates among dictators across norm treatment conditions.

Notes: Bars indicate 95% confidence intervals.

variation through a between-subjects design: participants are randomly assigned to one of several norm message conditions. I classify dictators into two groups based on their expectations about the majority norm behavior: *reveal expectants*, who believe that more than half (less than 50%) revealed the hidden information, and *ignorance expectants*, who believe that at half or less remain uninformed (50% or less revealed).[11] These terms reflect participants' beliefs about the prevailing descriptive norm—whether the majority seeks or avoids information.[12] Figure 3 shows the distribution of beliefs about reveal rates, which naturally separates at the 50% threshold. A Kruskal-Wallis test confirms that belief distributions do not differ significantly across the eight treatment arms ($\chi^2(7) = 6.45$, $p = 0.488$), supporting the comparability of subjects across conditions.

To explore how these beliefs correlate with actual behavior, Figure 4 plots average ignorance rates across bins of belief percentages. A clear negative trend emerges, with participants who believed more of their peers revealed information were themselves more likely to seek information, suggesting a behavioral anchoring effect consistent with norm

[11] As only 11 dictators predicted exactly 50% revealed—and only 4 of them chose to reveal—these participants most closely resembled *ignorance expectants* in behavior and were classified accordingly. Flipping this classification does not meaningful affect the results.

[12] Appendix K shows a demographics regression indicating no notable difference in the background of *ignorance expectants*.

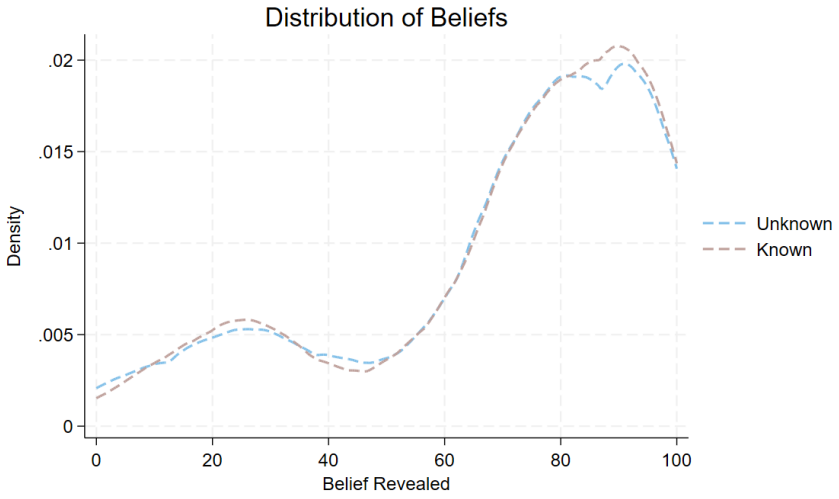


Figure 3: Distribution of participants' beliefs about the percentage of dictators who revealed the payoff tables.

sensitivity.

The difference in baseline behavior between the two groups is stark: under the *No Norm* condition, *ignorance expectant* dictators chose to remain ignorant 63% of the time, compared to just 22% among *reveal expectants*—a gap of over 40 percentage points. Figure 5 illustrates this gap by showing ignorance rates across treatment arms for both belief types. Appendix K explores whether demographic differences account for this pattern, while Appendix L examines differences in how each group rated the appropriateness of remaining ignorant.

The results suggest that individuals adjust their behavior based on how much their prior beliefs deviate from observed norms. This effect is particularly pronounced among *ignorance expectant* dictators—those who initially believed that most others would choose to remain ignorant. Although they represent a minority of the sample (N = 217), these individuals responded strongly when exposed to a *10% Ignorance* environment.[13] Specifically, for *ignorance expectant* dictators, exposure to the *10% Ignorance* environment decreased the likelihood of avoiding information by approximately 20 percentage points compared to the pooled sample of other categories. A two-sided difference

[13]In Appendix H, I check whether *ignorance expectant* dictators were more likely to exhibit norm-following characteristics by regressing the likelihood of being a *ignorance expectant* dictator on the subject's conformity score (Mehrabian and Stefl, 1995).

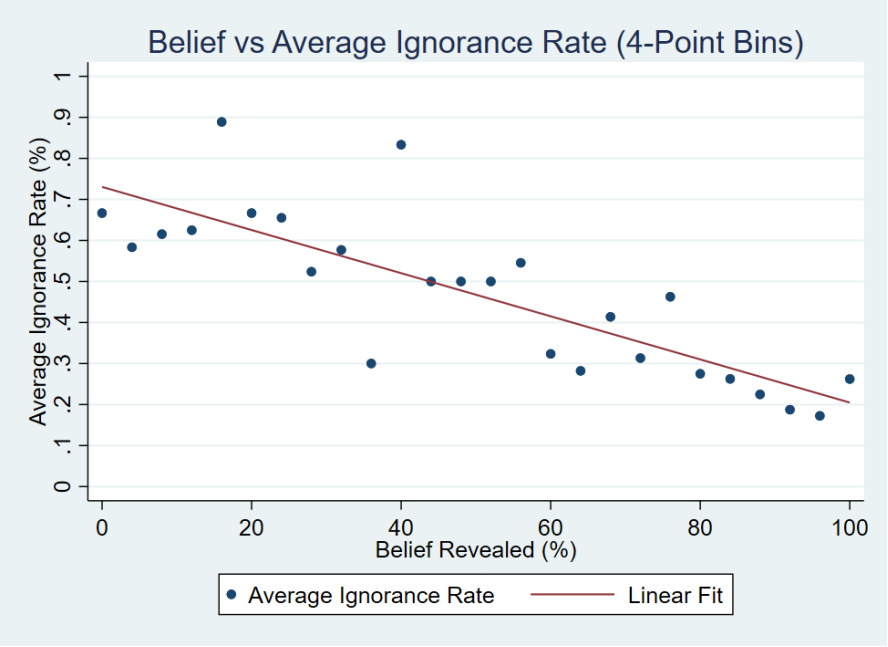
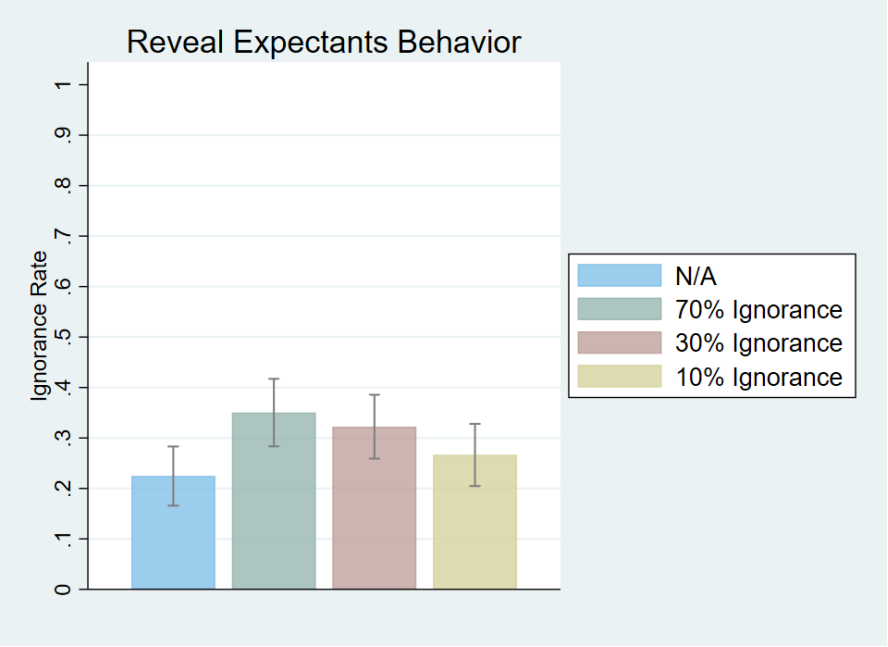


Figure 4: Average ignorance rate as a function of participants' beliefs about the percentage of dictators who revealed payoff tables.

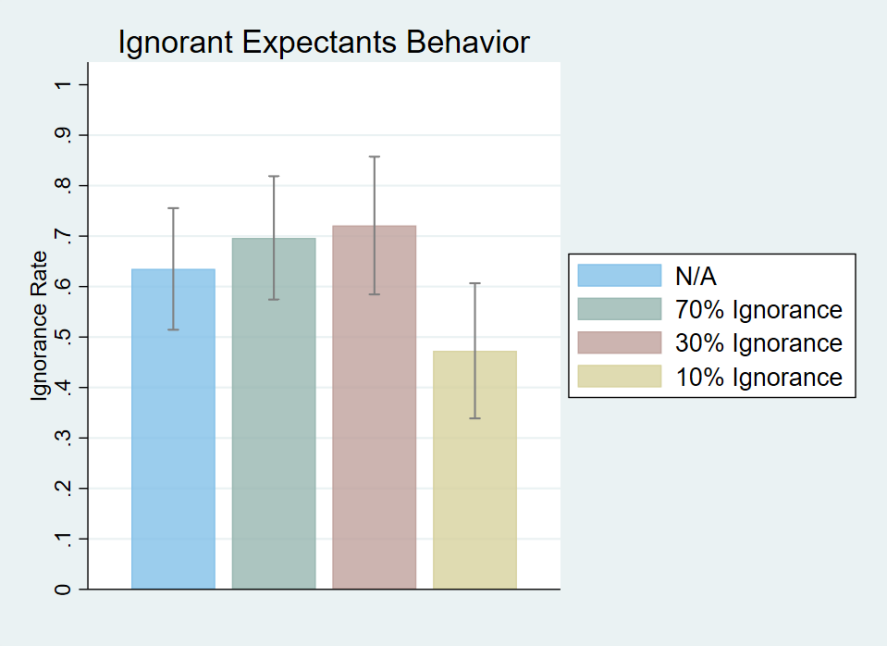
of means test confirms this effect is statistically significant ($t = 2.76$, $p < 0.001$). This provides clear evidence for *Hypothesis 2 - Norm Sensitivity*—the idea that individuals whose prior beliefs about ignorance deviate substantially from observed norms adjust their behavior accordingly.

Result 2: Compared to *reveal expectant* dictators, *ignorance expectant* dictators were far less likely to reveal payoffs but are significantly more likely to follow norms in the *10% Ignorance* environment.

In contrast, there is no support for *Hypothesis 3 - Exculpatory Norms*, which predicted that individuals whose prior beliefs understated the prevalence of ignorance (i.e., *reveal expectant* dictators) would be more likely to avoid information to justify selfish decisions. The data do not indicate that *reveal expectant* dictators strategically used ignorance as a justification. A complementary sentiment and linguistic analysis of participants' explanations supports this interpretation. As detailed in Appendix F, *reveal expectants* used more morally framed, positive language, while *ignorance expectants* employed instrumental reasoning that more resembles the behavior of conditional cooperators, whereby individuals adjust their behavior based on perceived norms or



(a) Ignorance rates of subjects who predicted an ignorance rate below 50%. N = 803.



(b) Ignorance rates of subjects who predicted an ignorance rate at or above 50%. N = 217.

Figure 5: Information avoidance rates among reveal expectant and ignorance expectant dictators across norm treatments.

Notes: Bars indicate 95% confidence intervals.

expectations about others (Fischbacher, Gächter, and Fehr, 2001). Instead, *reveal expectants'* information avoidance behavior remained largely unchanged regardless of the observed norm. These findings suggest that ignorance norms are not universally stable, but rather emerge and persist in environments where individuals hold similar prior expectations about others' ignorance. In such settings comprising primarily of *ignorance expectant* types, strategic avoidance of information can be mutually reinforced, leading to localized clusters of norm-consistent ignorance behavior (Momsen and Ohndorf, 2020).

Result 3: *Reveal expectant* dictators are not significantly more likely to adopt ignorance when exposed to a high ignorance environment.

To determine whether ex-ante self-serving beliefs are formed by dictators, I compare across the *Known* and *Unknown* conditions beliefs about perceived ignorance rates, elicited prior to the dictator allocation decision. A two-tailed difference of means test finds no significant differences in dictators' beliefs about the ignorance rates ($p = 0.75$). Furthermore, a Kolmogorov-Smirnov test indicate no significant difference in the distribution of reported beliefs ($D = 0.036, p = 0.897$). Figure 3 shows the distribution of these beliefs. The results suggest that dictators do not systematically adjust their beliefs about social norms in a self-serving manner before making their decisions, implying that other cognitive or social factors may drive the avoidance of information. This stands in contrast to prior findings on moral hypocrisy, which show that individuals often revise their fairness judgments after making selfish choices, suggesting a motivated attempt to resolve dissonance between self-image and behavior (Rustichini and Villeval, 2014).

Next, I examined the ex-post beliefs of dictators by analyzing both personal and social appropriateness ratings for choosing to reveal or not reveal the payoff tables, which were collected after participants made their allocation decisions. Social appropriateness ratings—elicited using coordination incentives to match the modal response—reflect second-order beliefs about how others perceive the action. In contrast, personal appropriateness ratings capture individuals' own moral judgments, which remained consistent across conditions. Comparing treatments, choosing not to reveal was rated as significantly more socially appropriate in the *70% Ignorance* condition, as confirmed by a Kruskal–Wallis test ($\chi^2(3) = 48.05$, $p < 0.001$), suggesting that injunctive norms update in response to descriptive ones. However, personal appropriateness ratings did not differ significantly across treatments for both choosing to reveal ($\chi^2(3) = 4.34$, $p = 0.228$) and choosing not to reveal ($\chi^2(3) = 4.08$, $p = 0.253$).

This divergence marks a clear departure from standard models of motivated cognition, which predict that individuals adjust personal beliefs or moral evaluations to justify self-interested behavior (Shepperd, Malone, and Sweeny, 2008). In contrast, the findings—further detailed in Appendix I—offer no support for Hypothesis 4 (Self-Serving Beliefs), suggesting that participants did not revise moral appropriateness ratings to align with their own choices. This reinforces the interpretation that individuals may compartmentalize personal norms and social expectations.[14] Alternatively, information avoidance may reflect fast, intuitive decision-making with limited concern for social judgment, consistent with dual-process models and recent evidence that such behavior persists even without image concerns (Exley and Kessler, 2023). A third interpretation is that eliciting beliefs about what others do may activate internal norms about acquiring information.[15]

Result 4: There is no evidence that dictators form self-serving beliefs.

These results suggest that the tendency to avoid information reflects a relatively stable personal preference for ignorance—one that persists even in the absence of explicit social cues. At the same time, behavior is clearly shaped by perceptions of what others typically do. Participants became more or less likely to acquire information depending on the norm environment they were exposed to, indicating that social expectations serve as an important behavioral anchor. This highlights a dual influence: a baseline individual inclination toward strategic ignorance and a social responsiveness to descriptive norms. The classification of participants into *ignorance expectants* and *reveal expectants* is indicative of the importance of underlying social types in shaping behavior. The absence of strong norm-induced shifts among most participants suggests that multiple behavioral equilibria are unlikely—except perhaps in contexts or groups dominated by *ignorance expectants*, where norm cues may exert stronger influence.[16]

4 Conclusion

This study demonstrates that individuals' willingness to avoid information is shaped by social norms, with lower information avoidance rates observed when norms favor information acquisition. Norm compliance differed between *ignorance expectants* who exhibited strong but selective norm responsiveness and *reveal expectants* who demon-

[14] Appendix L provides further heterogeneity analysis related to *ignorance expectants* behavior.
[15] Appendix N further discusses this interpretation.
[16] Appendix D presents a simple model of norm coordination supporting this intuition.

strated modest but consistent adjustments across norm environments. Importantly, behavior closely tracked norm messages, suggesting alignment with perceived expectations rather than belief distortion or motivated reasoning. However, this absence of belief distortion should not be taken to mean that people do not engage in self-serving cognition. Rather, it may reflect features of the experimental environment that constrained opportunities for such distortion. In settings with greater ambiguity or higher stakes, belief distortion may still emerge.

The experimental design implicitly treats information-seeking and information-avoidance as opposing normative social norms. That is, social environments can support either an information acquisition norm or an ignorance norm, and individuals' decisions may reflect attempts to coordinate with perceived expectations. Because descriptive norms were experimentally provided, participants had access to accurate information about others' behavior. The absence of belief distortion suggests that behavior was guided by genuine norm-following rather than motivated reasoning. This implies that, at least in settings with clearly communicated norms, social behavior may converge toward equilibrium through accurate norm perception.

While the shifts in information avoidance behavior in response to norm exposure are modest in absolute terms (8–12 percentage points), they are meaningful in the context of light-touch interventions. This translates to a relative reduction in ignorance behavior of roughly 20–30%, demonstrating that even subtle normative cues can produce sizable behavioral shifts in morally ambiguous settings. These findings point to scalable policy interventions where small nudges can shift behavior at scale, especially in settings where stronger interventions are impractical.

Notably, individuals' prior beliefs about others' behavior were broadly consistent with how people actually behaved, and there was no evidence that participants distorted these beliefs to justify their own decisions. This suggests that rather than engaging in motivated reasoning, individuals may simply view ignorance as an acceptable or normatively permissible option. As a result, shifting their behavior requires norm cues that are credible and meaningfully challenge prior expectations. The findings highlight that *ignorance expectants*—those who believe that ignorance is common—are particularly responsive to updated information about prevailing norms, but only when those norms are overwhelmingly one-sided. This responsiveness, however, may be especially relevant in environments where such individuals are clustered together. In these contexts, multiple ignorance norms may coexist and become self-reinforcing, sustained by shared expec-

tations and limited exposure to contradicting norm cues. Interventions targeting these groups with explicit messages about the widespread adoption of information-seeking behavior could therefore be particularly effective. For example, communications that highlight transparency as a majority behavior rather than an isolated practice may help dismantle ignorance belief clusters and encourage broader norm adherence.

Interestingly, ignorance rates were notably low even under the *No Norm* condition (32%) compared to Mol, Soraperra, and van der Weele (2025) and Z. Grossman, Hua, et al. (2025), who report rates above 60% using similar Prolific samples. One possible explanation is that elicitation of beliefs about norms made normative considerations more salient and influenced behavior. Such an effect could explain the lack of belief distortion in this study. Future research could test the behavioral effects of such elicitation directly. Another promising direction is to examine whether individuals strategically avoid learning not just about facts, but about the norms themselves. When normative expectations are ambiguous, acquiring norm information might impose stronger moral demands. In such cases, individuals may remain norm-ignorant to preserve moral flexibility, mirroring the logic of motivated ignorance in Spiekermann and Weiss (2016). Studying how people respond when norm information is endogenous could deepen our understanding of moral self-regulation.

Taken together, these findings suggest that strategic ignorance is not just self-serving convenience but a socially embedded behavior. People look to others for guidance in morally ambiguous situations, and when provided with accurate social cues, they adjust in ways that reflect real-time norm coordination.

AI Acknowledgement

During the preparation of this work, the author used OpenAI's ChatGPT and the Whitefull tool on Overleaf to revise the text for improved readability and flow. After using this service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication. As part of the analysis in Appendix F of this manuscript, ChatGPT was used to conduct sentiment analysis and linguistic interpretation on participants' open-ended responses. The tool was used to extract polarity and subjectivity scores from textual explanations using Python libraries (e.g., TextBlob).

References

Akerlof, George A. (1980). "A Theory of Social Custom, of Which Unemployment May be One Consequence". In: *The Quarterly Journal of Economics* 94.4, pages 749–775. ISSN: 00335533, 15314650. URL: http://www.jstor.org/stable/1885667.

Babcock, Linda and George Loewenstein (Mar. 1997). "Explaining Bargaining Impasse: The Role of Self-Serving Biases". In: *Journal of Economic Perspectives* 11.1, pages 109–126. DOI: 10.1257/jep.11.1.109. URL: https://www.aeaweb.org/articles?id=10.1257/jep.11.1.109.

Bašić, Zvonimir and Eugenio Verrina (2024). "Personal norms — and not only social norms — shape economic behavior". In: *Journal of Public Economics* 239, page 105255. ISSN: 0047-2727. DOI: https://doi.org/10.1016/j.jpubeco.2024.105255. URL: https://www.sciencedirect.com/science/article/pii/S0047272724001919.

Bicchieri, Cristina, Eugen Dimant, Michele Gelfand, and Silvia Sonderegger (2023). "Social norms and behavior change: The interdisciplinary research frontier". In: *Journal of Economic Behavior & Organization* 205, A4–A7. ISSN: 0167-2681. DOI: https://doi.org/10.1016/j.jebo.2022.11.007. URL: https://www.sciencedirect.com/science/article/pii/S0167268122004115.

Bicchieri, Cristina, Eugen Dimant, and Silvia Sonderegger (2023). "It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion". In: *Games and Economic Behavior* 138, pages 321–354. ISSN: 0899-8256. DOI: https://doi.org/10.1016/j.geb.2023.01.005. URL: https://www.sciencedirect.com/science/article/pii/S0899825623000088.

Charness, Gary and Matthew Rabin (2002). "Understanding Social Preferences with Simple Tests". In: *The Quarterly Journal of Economics* 117.3, pages 817–869. ISSN: 00335533, 15314650. URL: http://www.jstor.org/stable/4132490 (visited on 03/07/2025).

Croson, Rachel and Uri Gneezy (June 2009). "Gender Differences in Preferences". In: *Journal of Economic Literature* 47.2, pages 448–74. DOI: 10.1257/jel.47.2.448. URL: https://www.aeaweb.org/articles?id=10.1257/jel.47.2.448.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1, pages 67–80. ISSN: 09382259, 14320479. URL: http://www.jstor.org/stable/27822583.

Eckel, Catherine C. and Philip J. Grossman (1998). "Are Women Less Selfish Than Men?: Evidence From Dictator Experiments". In: *The Economic Journal* 108.448, pages 726–735. DOI: https://doi.org/10.1111/1468-0297.00311. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0297.00311. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0297.00311.

Enke, Benjamin, Raymond Fisman, Luis Mota Freitas, and Steven Sun (June 2024). "Universalism and Political Representation: Evidence from the Field". In: *American Economic Review: Insights* 6.2, pages 214–29. DOI: 10.1257/aeri.20230222. URL: https://www.aeaweb.org/articles?id=10.1257/aeri.20230222.

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann (2022). "Moral Universalism: Measurement and Economic Relevance". In: *Management Science* 68.5, pages 3590–3603. DOI: 10.1287/mnsc.2021.4086. eprint: https://doi.org/10.1287/mnsc.2021.4086. URL: https://doi.org/10.1287/mnsc.2021.4086.

Exley, Christine L. and Judd B. Kessler (Aug. 2023). "Information Avoidance and Image Concerns". In: *The Economic Journal* 133.656, pages 3153–3168. ISSN: 0013-0133. DOI: 10.1093/ej/uead058. eprint: https://academic.oup.com/ej/article-pdf/133/656/3153/51918010/uead058.pdf. URL: https://doi.org/10.1093/ej/uead058.

Exley, Christine L. and Judd B. Kessler (Apr. 2024). "Motivated Errors". In: *American Economic Review* 114.4, pages 961–87. DOI: 10.1257/aer.20191849. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20191849.

Feiler, Lauren (2014). "Testing models of information avoidance with binary choice dictator games". In: *Journal of Economic Psychology* 45, pages 253–267. ISSN: 0167-4870. DOI: https://doi.org/10.1016/j.joep.2014.10.003. URL: https://www.sciencedirect.com/science/article/pii/S0167487014000804.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001). "Are people conditionally cooperative? Evidence from a public goods experiment". In: *Economics Letters* 71.3, pages 397–404. ISSN: 0165-1765. DOI: https://doi.org/10.1016/S0165-1765(01)00394-9. URL: https://www.sciencedirect.com/science/article/pii/S0165176501003949.

Giamattei, Marcus, Kaveh S. Yahosseini, Simon Gächter, and Lucas Molleman (2020). "LIONESS Lab: a Free Web-Based Platform for Conducting Interactive Experiments Online". In: *Journal of the Economic Science Association* 6, pages 95–111. DOI: 10.1007/s40881-020-00087-0.

Gino, Francesca, Michael I. Norton, and Roberto A. Weber (Sept. 2016). "Motivated Bayesians: Feeling Moral While Acting Egoistically". In: *Journal of Economic Per-*

spectives 30.3, pages 189–212. DOI: 10.1257/jep.30.3.189. URL: https://www.aeaweb.org/articles?id=10.1257/jep.30.3.189.

Gneezy, Uri, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen (2020). "Bribing the Self". In: *Games and Economic Behavior* 120, pages 311–324. ISSN: 0899-8256. DOI: https://doi.org/10.1016/j.geb.2019.12.010. URL: https://www.sciencedirect.com/science/article/pii/S0899825619301939.

Golman, Russell, David Hagmann, and George Loewenstein (Mar. 2017). "Information Avoidance". In: *Journal of Economic Literature* 55.1, pages 96–135. DOI: 10.1257/jel.20151245. URL: https://www.aeaweb.org/articles?id=10.1257/jel.20151245.

Grossman, Zachary (2014). "Strategic ignorance and the robustness of social preferences". In: *Management Science* 60.11, pages 2659–2665. ISSN: 00251909, 15265501. URL: http://www.jstor.org/stable/24550536.

Grossman, Zachary, Tony Hua, Jo Thori Lind, and Karine Nyborg (Jan. 2025). *Unwillingly Informed: the Prosocial Impact of Third-Party Informers*. Memorandum 2/2025. Oslo University, Department of Economics. URL: https://ideas.repec.org/p/hhs/osloec/2025_002.html.

Grossman, Zachary and Joël J. Van der Weele (Dec. 2017). "Self-image and willful ignorance in social decisions". In: *Journal of the European Economic Association* 15.1, pages 173–217. ISSN: 1542-4766. DOI: 10.1093/jeea/jvw001. eprint: https://academic.oup.com/jeea/article-pdf/15/1/173/11079001/jvw001.pdf. URL: https://doi.org/10.1093/jeea/jvw001.

Krupka, Erin L. and Roberto A. Weber (June 2013). "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?" In: *Journal of the European Economic Association* 11.3, pages 495–524. ISSN: 1542-4766. DOI: 10.1111/jeea.12006. eprint: https://academic.oup.com/jeea/article-pdf/11/3/495/10317174/jeea0495.pdf. URL: https://doi.org/10.1111/jeea.12006.

Lind, Jo Thori, Karine Nyborg, and Anna Pauls (2019). "Save the planet or close your eyes? Testing strategic ignorance in a charity context". In: *Ecological Economics* 161, pages 9–19. ISSN: 0921-8009. DOI: https://doi.org/10.1016/j.ecolecon.2019.02.010. URL: https://www.sciencedirect.com/science/article/pii/S0921800918312667.

Mehrabian, Albert and Carol A. Stefl (1995). "Basic temperament components of loneliness, shyness, and conformity". In: *Social Behavior and Personality: An International Journal* 23.3, pages 253–263. DOI: 10.2224/sbp.1995.23.3.253.

Mol, Jantsje M., Ivan Soraperra, and Joël J. van der Weele (2025). "Spoiling the party: Experimental evidence on the willingness to transmit inconvenient ethical information". In: *Experimental Economics*, pages 1–19. DOI: 10.1017/eec.2025.6.

Momsen, Katharina and Markus Ohndorf (2020). "When do people exploit moral wiggle room? An experimental analysis of information avoidance in a market setup". In: *Ecological Economics* 169, page 106479. ISSN: 0921-8009. DOI: https://doi.org/10.1016/j.ecolecon.2019.106479. URL: https://www.sciencedirect.com/science/article/pii/S0921800919303738.

Momsen, Katharina and Markus Ohndorf (Apr. 2023). "Information avoidance: Self-image concerns, inattention, and ideology". In: *Journal of Economic Behavior & Organization* 211, pages 386–400. DOI: 10.1016/j.jebo.2023.04.032.

Rustichini, Aldo and Marie Claire Villeval (2014). "Moral hypocrisy, power and social preferences". In: *Journal of Economic Behavior & Organization* 107, pages 10–24. ISSN: 0167-2681. DOI: https://doi.org/10.1016/j.jebo.2014.08.002. URL: https://www.sciencedirect.com/science/article/pii/S0167268114002169.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius (2007). "The Constructive, Destructive, and Reconstructive Power of Social Norms". In: *Psychological Science* 18.5. PMID: 17576283, pages 429–434. DOI: 10.1111/j.1467-9280.2007.01917.x. eprint: https://doi.org/10.1111/j.1467-9280.2007.01917.x. URL: https://doi.org/10.1111/j.1467-9280.2007.01917.x.

Shepperd, James, Wendi Malone, and Kate Sweeny (2008). "Exploring Causes of the Self-serving Bias". In: *Social and Personality Psychology Compass* 2.2, pages 895–908. DOI: https://doi.org/10.1111/j.1751-9004.2008.00078.x. eprint: https://compass.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-9004.2008.00078.x. URL: https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-9004.2008.00078.x.

Spiekermann, Kai and Arne Weiss (2016). "Objective and subjective compliance: A norm-based explanation of 'moral wiggle room'". In: *Games and Economic Behavior* 96, pages 170–183. ISSN: 0899-8256. DOI: https://doi.org/10.1016/j.geb.2015.11.007. URL: https://www.sciencedirect.com/science/article/pii/S0899825615001554.

Tomasello, Michael, Ann C. Kruger, and Hilary H. Ratner (1993). "Cultural learning". In: *Behavioral and Brain Sciences* 16.3, pages 495–552. DOI: 10.1017/S0140525X0003123X.

van der Weele, Joël J., Julija Kulisa, Michael Kosfeld, and Guido Friebel (Aug. 2014). "Resisting Moral Wiggle Room: How Robust Is Reciprocal Behavior?" In: *American Economic Journal: Microeconomics* 6.3, pages 256–64. DOI: 10.1257/mic.6.3.256. URL: https://www.aeaweb.org/articles?id=10.1257/mic.6.3.256.

Vu, Lam, Ivan Soraperra, Margarita Leib, Joel van der Weele, and Shaul Shalvi (2023). "Ignorance by choice: A meta-analytic review of the underlying motives of willful ignorance and its consequences". In: *Psychological Bulletin* 149.9-10, pages 611–635. DOI: 10.1037/bul0000398. URL: https://doi.org/10.1037/bul0000398.

A Full Demographics

This section provides an overview of the demographic characteristics of the study participants across different experimental conditions. The sample includes 1,020 participants recruited via Prolific, with a balanced distribution across age, gender, education, and political affiliation. Table 3 reports the full breakdown of subject's demographics.

Most participants are in their late 30s, with a roughly equal gender split. The majority have at least a high school education, and political affiliations are fairly evenly distributed among Democrats, Independents, and Republicans. Religious backgrounds vary, with Christianity being the most common.

The demographic balance across conditions ensures that the study's findings are not driven by sample differences. Further details and statistical tests confirming this balance are reported in Appendix J.

Table 3: This table presents detailed demographic characteristics of participants across all treatment conditions. Variables include age, gender, student status, race/ethnicity, income, education, political affiliation, and religious identity. Each cell shows the proportion or mean value for the group, with standard deviations in parentheses where applicable.

	Treatment Arms								
	Unk. N/A n=132	Unk. 70% Ig. n=119	Unk. 30% Ig. n=119	Unk. 10% Ig. n=134	Kwn. N/A n=127	Kwn. 70% Ig. n=134	Kwn. 30% Ig. n=135	Kwn. 10% Ig. n=120	Total n=1,020
age	38.30 (12.10)	38.27 (12.48)	38.80 (11.10)	39.02 (11.44)	39.11 (13.67)	37.12 (12.73)	37.33 (12.51)	38.13 (11.60)	38.25 (12.22)
female	0.55 (0.50)	0.51 (0.50)	0.58 (0.50)	0.51 (0.50)	0.52 (0.50)	0.42 (0.50)	0.46 (0.50)	0.44 (0.50)	0.50 (0.50)
student	0.17 (0.38)	0.17 (0.37)	0.06 (0.24)	0.10 (0.30)	0.12 (0.33)	0.16 (0.37)	0.12 (0.33)	0.12 (0.33)	0.13 (0.33)
race									
White	94 (74.6%)	80 (73.4%)	77 (69.4%)	82 (66.1%)	89 (71.2%)	90 (70.9%)	94 (74.6%)	88 (76.5%)	694 (72.1%)
Black	12 (9.5%)	12 (11.0%)	13 (11.7%)	18 (14.5%)	10 (8.0%)	21 (16.5%)	13 (10.3%)	11 (9.6%)	110 (11.4%)
Hispanic	8 (6.3%)	5 (4.6%)	10 (9.0%)	9 (7.3%)	6 (4.8%)	8 (6.3%)	7 (5.6%)	2 (1.7%)	55 (5.7%)
East Asian	7 (5.6%)	4 (3.7%)	3 (2.7%)	4 (3.2%)	8 (6.4%)	4 (3.1%)	3 (2.4%)	5 (4.3%)	38 (3.9%)
Southeast Asian	3 (2.4%)	5 (4.6%)	4 (3.6%)	4 (3.2%)	8 (6.4%)	3 (2.4%)	3 (2.4%)	4 (3.5%)	34 (3.5%)
MENA	1 (0.8%)	1 (0.9%)	0 (0.0%)	2 (1.6%)	1 (0.8%)	0 (0.0%)	1 (0.8%)	0 (0.0%)	6 (0.6%)
Other Race	1 (0.8%)	2 (1.8%)	4 (3.6%)	5 (4.0%)	3 (2.4%)	1 (0.8%)	5 (4.0%)	5 (4.3%)	26 (2.7%)
income									
$0	6 (4.9%)	4 (3.7%)	1 (0.9%)	4 (3.6%)	6 (5.1%)	2 (1.7%)	5 (4.3%)	2 (1.8%)	30 (3.3%)
less than $20,000	23 (18.9%)	23 (21.5%)	15 (14.0%)	21 (18.8%)	27 (23.1%)	30 (26.1%)	26 (22.2%)	30 (27.5%)	195 (21.5%)
less than $40,000	14 (11.5%)	10 (9.3%)	23 (21.5%)	20 (17.9%)	19 (16.2%)	19 (16.5%)	27 (23.1%)	16 (14.7%)	148 (16.3%)
less than $60,000	23 (18.9%)	17 (15.9%)	18 (16.8%)	20 (17.9%)	25 (21.4%)	23 (20.0%)	22 (18.8%)	16 (14.7%)	164 (18.1%)
less than $80,000	23 (18.9%)	22 (20.6%)	18 (16.8%)	14 (12.5%)	19 (16.2%)	15 (13.0%)	16 (13.7%)	18 (16.5%)	145 (16.0%)
less than $100,000	11 (9.0%)	12 (11.2%)	11 (10.3%)	11 (9.8%)	7 (6.0%)	8 (7.0%)	9 (7.7%)	11 (10.1%)	80 (8.8%)
greater than $100,000	22 (18.0%)	19 (17.8%)	21 (19.6%)	22 (19.6%)	14 (12.0%)	18 (15.7%)	12 (10.3%)	16 (14.7%)	144 (15.9%)
education									
Less than high school	0 (0.0%)	0 (0.0%)	1 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.8%)	0 (0.0%)	2 (0.2%)
High school graduate	41 (33.6%)	45 (42.5%)	40 (37.7%)	40 (35.7%)	50 (42.7%)	47 (40.9%)	55 (46.6%)	37 (33.9%)	355 (39.2%)
Bachelors	61 (50.0%)	41 (38.7%)	43 (40.6%)	53 (47.3%)	45 (38.5%)	56 (48.7%)	48 (40.7%)	52 (47.7%)	399 (44.1%)
Masters	16 (13.1%)	18 (17.0%)	22 (20.8%)	19 (17.0%)	22 (18.8%)	8 (7.0%)	11 (9.3%)	17 (15.6%)	133 (14.7%)
Doctorate	4 (3.3%)	2 (1.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (3.5%)	3 (2.5%)	3 (2.8%)	16 (1.8%)
political_party									
Republican	34 (27.9%)	22 (20.8%)	24 (22.9%)	26 (23.2%)	24 (20.7%)	26 (22.6%)	27 (22.9%)	27 (24.8%)	210 (23.3%)
Independent	41 (33.6%)	42 (39.6%)	41 (39.0%)	39 (34.8%)	40 (34.5%)	42 (36.5%)	42 (35.6%)	29 (26.6%)	316 (35.0%)
Democrat	47 (38.5%)	42 (39.6%)	40 (38.1%)	47 (42.0%)	52 (44.8%)	47 (40.9%)	49 (41.5%)	53 (48.6%)	377 (41.7%)
religion									
Christianity	58 (47.9%)	48 (45.3%)	54 (51.4%)	52 (46.4%)	50 (43.5%)	57 (49.6%)	51 (43.2%)	50 (45.9%)	420 (46.6%)
Islam	1 (0.8%)	2 (1.9%)	1 (1.0%)	3 (2.7%)	1 (0.9%)	0 (0.0%)	2 (1.7%)	0 (0.0%)	10 (1.1%)
Judaism	0 (0.0%)	4 (3.8%)	2 (1.9%)	5 (4.5%)	3 (2.6%)	2 (1.7%)	2 (1.7%)	1 (0.9%)	19 (2.1%)
Hinduism	0 (0.0%)	0 (0.0%)	1 (1.0%)	0 (0.0%)	2 (1.7%)	0 (0.0%)	1 (0.8%)	1 (0.9%)	5 (0.6%)
Buddhism	2 (1.7%)	1 (0.9%)	0 (0.0%)	1 (0.9%)	5 (4.3%)	0 (0.0%)	1 (0.8%)	1 (0.9%)	11 (1.2%)
Agnosticism	20 (16.5%)	23 (21.7%)	20 (19.0%)	20 (17.9%)	22 (19.1%)	27 (23.5%)	29 (24.6%)	24 (22.0%)	185 (20.5%)
Atheism	21 (17.4%)	20 (18.9%)	15 (14.3%)	14 (12.5%)	16 (13.9%)	21 (18.3%)	17 (14.4%)	20 (18.3%)	144 (16.0%)
Other Religion	19 (15.7%)	8 (7.5%)	12 (11.4%)	17 (15.2%)	16 (13.9%)	8 (7.0%)	15 (12.7%)	12 (11.0%)	107 (11.9%)

B Pilot Studies

Two pilot studies were conducted at the University of California, Merced. The first study, involving 51 subjects, took place in November 2024. It included only the *Known - No Info* and *Unknown - No Info* treatment arms and was not incentivized. The second pilot study, conducted in January 2025 with 75 subjects, introduced additional treatments, incorporating *Known/Unknown* and *No Info/70%/30% Ignorance* conditions. To partially incentivize participation, four subjects were randomly selected to receive payment based on their decisions.

Belief distributions in the pilot studies mirrored those observed in the main experiment, with participants clustering into two distinct types: *reveal expectants* and *ignorance expectants*. As most subjects in the pilot studies predicted an ignorance rate lower than 30%, an additional *10% Ignorance* treatment arm was added to the main experiment. This allowed for a sufficient sample size to assess the behavior of subjects who underestimated the reveal rate. Several textual refinements were made to the experimental interface between the pilot and main studies. Due to changes in instructions, the lack of incentives in the first pilot, and differences in the subject pool, direct comparisons between the pilot and main study results should be interpreted with caution.

In the first and second pilot studies, the average ignorance rates were 55% and 45%, respectively. Across both studies, subjects in the *Known* condition exhibited a 10-12% higher ignorance rate compared to those in the *Unknown* condition. Notably, participants in the *Known - No Info* condition acquired payoff information at a significantly lower rate (43%) compared to those in the *Unknown - No Info* condition (56%). This 13-percentage point difference suggests that making one's decision-making role salient from the outset may increase the likelihood of strategic ignorance. One possible explanation is heightened awareness of self-serving incentives; another is the effect of *self-referential* framing—participants in the *Known* condition were addressed as "you," whereas those in the *Unknown* condition evaluated behavior attributed to "Person 1." Given the weak or absent incentivization in these pilot studies, the latter explanation—framing effects—appears more likely.

In addition to demographic differences between the pilot studies and the main Prolific sample, the testing environment may have influenced behavior. Pilot participants were seated in-person, potentially fostering a more socially conscious or norm-sensitive atmosphere. In contrast, Prolific subjects completed the study online in private settings. Demographically, the Prolific sample skewed older, predominantly white, and

largely non-student, while the pilot participants were primarily younger university students from more racially diverse backgrounds.

C Role Relevation and Self-Referential Framinig

An unintentional but consequential artifact of the experimental design arose from the timing of when participants learned they would be making the allocation decision themselves. In the *Known* condition, subjects were told from the outset that they would act as the decision-maker, whereas in the *Unknown* condition, they were initially asked to reason about what another person ("Person 1") would do before later learning that they would assume that role. While this variation was originally intended to test how belief formation differs when individuals are primed to reason from a personal versus third-party perspective, subsequent analysis suggests that it also introduced a subtle but powerful framing effect. Specifically, because both conditions ultimately involved participants making the same decision with the same information, the difference cannot be explained by role knowledge alone. Instead, behavioral patterns from lightly incentivized pilot studies—which also showed effects consistent with the *self-referential* versus *socially contextualized* distinction—suggest that it is the framing of agency and identity ("you" vs. "Person 1"), rather than the timing of role revelation, that drives these differences. This realization motivates closer examination of how subtle linguistic and narrative cues may shape moral decision-making, particularly in norm-sensitive environments.

The downstream implications of this framing artifact are explored in Appendix G, which examines its influence on reported moral universalism scores, and in Appendix M, which shows that women were less likely than men to follow descriptive norms when decisions were framed *self-referentially*.

D Theoretical Framework

A Utility Model of Norm-Sensitive Ignorance Behavior

To interpret the observed asymmetries in how individuals respond to norm cues, I develop a simple utility model of strategic ignorance with belief-type-dependent norm sensitivity. The model builds on the logic of conditional norm-following Bicchieri, Dimant, Gelfand, et al. (2023), in which individuals align their behavior with perceived de-

scriptive norms when they believe those norms are both commonly followed and socially expected. Following Spiekermann and Weiss (2016), I formalize this process as a form of norm-based equilibrium coordination, where individuals derive utility not only from material payoffs but also from conforming to empirical expectations. This framework helps explain why *ignorance expectants*—individuals who initially believe that most others avoid information—respond strongly to pro-information norms that sharply contradict their expectations, while *reveal expectants* exhibit more gradual adjustment across all norm conditions.

Let the individual choose whether to acquire information about the recipient's payoff ($I = 1$) or to remain ignorant ($I = 0$). This decision precedes an allocation choice, which determines the material payoff $a \in \mathbb{R}$. Individuals may also experience a psychological cost or benefit from conforming to—or deviating from—perceived descriptive norms.

The following parameters describe the model:

- $\mu \in [0, 1]$: perceived descriptive norm, i.e., proportion of others who choose to reveal
- $b \in [0, 1]$: individual's prior belief about others' behavior (elicited before feedback)
- $\eta(\mu, b) \geq 0$: norm conformity weight, varying by belief type and belief-norm misalignment

The utility of each action is given by:

$$U_1 = a_1 + \eta(\mu, b) \cdot (2\mu - 1)$$
$$U_0 = a_0 - \eta(\mu, b) \cdot (2\mu - 1)$$

where:

- U_1: utility from revealing the payoff information
- U_0: utility from avoiding the information
- $(2\mu - 1) \in [-1, 1]$: direction and strength of the perceived norm

To account for the empirical asymmetry in norm responsiveness, the norm conformity weight $\eta(\mu, b)$ is defined piecewise:

$$\eta(\mu, b) = \begin{cases} \eta_{IE} & \text{if } b < 0.5 \text{ and } \mu < 0.3 \\ \eta_{RE} & \text{if } b > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where:

- $\eta_{IE} > \eta_{RE} > 0$
- $b < 0.5$: *Ignorance Expectant* (expects most others will remain ignorant)
- $b > 0.5$: *Reveal Expectant* (expects most others will reveal)

This formulation implies that:

(a) ***Ignorance expectants*** respond strongly to sharply contradicting norms—particularly when exposed to a *pro-information* norm (e.g., $\mu = 0.1$), which challenges their belief that ignorance is socially typical.

(b) ***Reveal expectants*** exhibit modest conformity across norm treatments but do not require strong contradiction to adjust.

(c) **No norm pressure arises** when the cue confirms prior beliefs or is too ambiguous to elicit a normative shift.

This model captures the core behavioral finding from the experiment: norm-induced behavior change depends not only on the strength of the norm cue but also on whether it violates the individual's prior expectations. It explains why only *ignorance expectants*—who assume ignorance is widespread—reduce information avoidance only when exposed to a strong pro-reveal signal. *Reveal expectants*, in contrast, show muted adjustments across all cue conditions.

By allowing the norm conformity term to be belief-type-dependent and zero in cases of norm alignment or ambiguity, the model formalizes the cognitive boundary conditions for norm-sensitive strategic ignorance.

Pooling Norms and Endogenous Equilibria

While the main model treats the descriptive norm μ as exogenously given (e.g., through norm feedback in the experiment), it is also possible to endogenize norm formation by treating μ as a belief about others' behavior. That is, individuals expect others to act according to the same behavioral rule they themselves follow, leading to the possibility of a self-fulfilling social norm.

Let μ^* denote the equilibrium level of information acquisition in the population. In equilibrium, μ^* must equal the expected proportion of individuals who choose to reveal information, given their beliefs and preferences. Assuming a logit choice framework, where individuals choose to reveal with probability increasing in the utility difference between actions, we define:

$$\mu^* = \mathbb{E}_b \left[\frac{\exp(U_1(b, \mu^*))}{\exp(U_1(b, \mu^*)) + \exp(U_0(b, \mu^*))} \right]$$

where $U_1(b, \mu)$ and $U_0(b, \mu)$ are the utility values from revealing and avoiding information, respectively, as defined in the base model. The expectation is taken over the population distribution of belief types b (e.g., a uniform or empirically estimated distribution).

This fixed-point equation expresses that in equilibrium, individuals' belief about the norm (i.e., the expected rate at which others reveal information) must match the actual aggregate behavior generated by those same expectations.

Multiple Equilibria. Because norm sensitivity $\eta(\mu, b)$ varies with prior belief b, and because conformity behavior may sharply increase when norms contradict expectations, this model can admit multiple equilibria. For example:

- A *low-μ^** equilibrium, in which most individuals avoid information and believe that others do too (*ignorance pooling*).
- A *high-μ^** equilibrium, in which most individuals acquire information and expect information acquisition to be the norm (*informed pooling*).

Such multiplicity is consistent with the experimental observation that *ignorance expectants* are particularly responsive to pro-information cues, while *reveal expectants* exhibit more stable behavior across norm conditions. The existence of stable ignorance clusters may therefore reflect coordination on a low-μ^* equilibrium, sustained by shared expectations and asymmetric norm responsiveness.

This extension highlights the potential for normative environments to lock in divergent behavioral patterns, even in otherwise identical decision contexts, and suggests that strategic ignorance may be more persistent in groups with converging priors or limited norm feedback.

E Robustness Check Regression on Ignorance Norms

As a robustness check for Result 1, I estimate a probit regression where the dependent variable is an indicator for ignorance (1 = did not reveal the payoff table). The key independent variable is a numeric treatment measure coded as 70, 30, or 10, reflecting

the actual rates of ignorance observed in each norm condition (i.e., the percentage of prior participants who chose to remain ignorant).

The model includes demographic controls: age, gender, race, income, education, employment status, student status, political affiliation, and religion. I find a positive and statistically significant effect, indicating that higher ignorance environments increases the likelihood of information avoidance. The results confirm a monotonic decline in ignorance as norms increasingly favor information acquisition.

Table 4: Probit Regression on Ignorance Choice and Ignorance Environments

	(1) avoided info	(2) avoided info
avoided info		
norm_ignorance_rate	0.00471**	0.00479**
	(0.00186)	(0.00209)
Controls	No	Yes
Observations	761	655

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

F Sentiment and Linguistic Analysis of Subjects' Open Response

To assist in interpreting the sentiment and thematic patterns, I used OpenAI's ChatGPT (version GPT-4, accessed July 2025) to synthesize trends across groups. ChatGPT was used exclusively for descriptive synthesis and interpretation; no generated text was used verbatim in the manuscript. All underlying data were produced by participants as part of the experiment. No model fine-tuning or custom training was applied. Using Python's TextBlob library, I extracted sentiment polarity (valence from –1 to +1) and subjectivity (scale from 0 to 1) for each explanation. Additionally, I used TF-IDF (term frequency–inverse document frequency) to identify prominent words in each group's reasoning. The results are presented in Table 5.

Reveal expectants expressed higher polarity (mean = 0.214) and greater subjectivity

(mean = 0.503), suggesting more positive and opinion-driven explanations. Their most common terms—such as "right," "fair," "know," "others," and "information"—reflect moral language, indicating that many framed information acquisition as a normative good. One such participant wrote, *"Since it's free and helps everyone make a better decision, I figured most would choose to reveal it to keep things fair and transparent,"* signaling a belief in shared prosocial values. This pattern aligns with the behavioral finding that *reveal expectants* were relatively insensitive to descriptive norm treatments, consistent with internally anchored moral motivation.

By contrast, *ignorance expectants* used more instrumental and self-interested language, marked by lower polarity (0.159) and subjectivity (0.456). Their dominant terms—"maximize," "money," "gain," "avoid," and "easy"—reflect a reasoning style focused on utility maximization, cost-benefit tradeoffs, and strategic simplicity. For example, one *ignorance expectant* explained, *"I thought that the majority of people would look to maximize their own earnings without caring what Person 2 received,"* emphasizing assumptions about others' self-interest rather than moral considerations.

This linguistic framing is consistent with the behavioral pattern that *ignorance expectants* were significantly more responsive to norm information, especially in the *10% Ignorance* condition. In one interpretation, *ignorance expectants* resemble conditional cooperators—individuals who are not intrinsically moralizers but who adjust their behavior when they perceive prosocial action as common or expected. Their reasoning and behavior reflect strategic norm-following rather than principled commitment, suggesting that norm-based interventions may be particularly effective within this subgroup.

Together, these results reinforce the paper's main conclusion. While some individuals exhibit stable moral motivation, others engage in strategic ignorance based on perceived normative expectations. Recognizing these differences is essential for designing effective nudges or norm-based interventions in ethical decision-making contexts.

G Moral Universalism Trust Survey

At the end of the questionnaire, participants had the option to complete the Moral Universalism Trust Survey, which measures the extent of trust individuals place in people across different social distances (Enke, Rodríguez-Padilla, and Zimmermann, 2022). A higher score reflects greater generalized trust toward strangers, whereas a lower score indicates a stronger preference for in-group trust.

Table 5: Sentiment and Linguistic Theme Analysis by Expectant Type

Measure	**Reveal Expectants**	**Ignorance Expectants**	**Interpretation**
Sentiment Polarity (mean ± SD)	0.214 ± 0.232	0.159 ± 0.237	*Reveal expectants* used more positive tone
Subjectivity (mean ± SD)	0.503 ± 0.230	0.456 ± 0.253	*Reveal expectants* expressed more opinion-driven reasoning
Top Linguistic Themes (TF-IDF)	*right, fair, know, others, information*	*maximize, money, gain, easy, avoid*	Moral framing vs. instrumental reasoning
Moral Reasoning Style	Deontological Principled	Conditional Norm-dependent	Reflects differences in norm sensitivity
Motivational Profile	Internalized norms	Strategic norm-following	Matches observed behavioral patterns

The regression table below examines the relationship between a participant's universalism score and their likelihood of revealing the payoff tables. Interestingly, universalism is negatively correlated with information avoidance behavior only in the *Known* treatments, suggesting that a *self-referential* framing engages one's universalism traits.

Table 6: Effect of Universalism Score on the Likelihood of Avoiding Information

	(1) avoided info	(2) avoided info	(3) avoided info	(4) avoided info
universalism score	-0.00742*	-0.00977**	-0.00280	-0.00327
	(0.00393)	(0.00480)	(0.00411)	(0.00499)
Condition	Known	Known	Unknown	Unknown
Controls	No	Yes	No	Yes
Observations	341	328	296	291

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The distribution of subjects' universalism trust scores does not significantly differ between the two groups. A Kolmogorov-Smirnov test indicates that the distributions are statistically indistinguishable ($D = 0.0662, p = 0.491$). Figure 6 shows the distribution of universalism scores.

As additional collaborative evidence, Table 9 shows that in the *Unknown* condition,

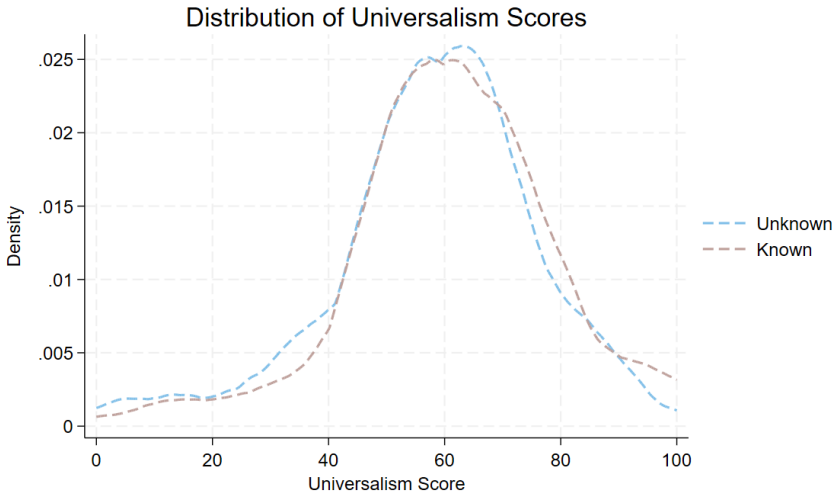


Figure 6: Distribution of participants' Universalism Trust scores.

political affiliation has no statistical effect, while in the *Known* condition, Republications, relative to Democrats, were more likely to avoid information, a result closely aligned with Momsen and Ohndorf (2023). Furthermore, these results are in line with Enke, Fisman, et al. (2024) who found that universalism is positively correlated with liberal beliefs. Given that universalism becomes more statistically significant once controls are added, I consider this suggestive evidence that *self-centric* framing activates innate traits.

H Conformity Score

In this appendix, I examine the Conformity Score, a psychological measure of norm-following behavior (Mehrabian and Stefl, 1995). The first step is to assess whether this score has predictive power. To do this, I create a binary variable, followed_norm, which equals 1 if a subject follows the majority's behavior—choosing to reveal when the majority revealed or choosing not to reveal when the majority did not reveal. The regression analysis suggests a positive correlation between the Conformity Score and norm-following behavior, but the results are not statistically significant ($p = 0.19$). Results reported in Table 7.

Next, I investigate whether *ignorance expectant* dictators are more likely to conform to norms. To test this, I regress subjects' Conformity Scores—used as a proxy for norm-

following tendencies—against their likelihood of being classified as *ignorance expectant* dictators. The results indicate no statistically significant relationship, suggesting that neither *ignorance expectant* nor *reveal expectant* dictators are inherently more likely classified as norm followers.

Table 7: Effect of Conformity Score on the Likelihood of Following Norms or Being a *Ignorance Expectant* Dictator

	(1) followed_norm	(2) followed_norm	(3) Ignorance Expectant	(4) Ignorance Expectant
main				
conformity score	0.0548	0.0611	-0.00387	-0.0104
	(0.0421)	(0.0464)	(0.0477)	(0.0518)
Controls	No	Yes	No	Yes
Observations	675	654	675	637

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I Social Appropriateness Ratings of Ignorance

This appendix examines self-reported social appropriateness ratings associated with the decision not to reveal information. Two distinct measures of social appropriateness were collected:

1. Social Appropriateness: Participants predicted the modal response of other subjects regarding the appropriateness of not revealing. They were incentivized with an additional ECU for correctly matching the most common response.

2. Personal Appropriateness: Participants provided their own personal evaluation of whether choosing not to reveal was appropriate, uninfluenced by incentive alignment.

The following pre-registered hypothesis was tested:

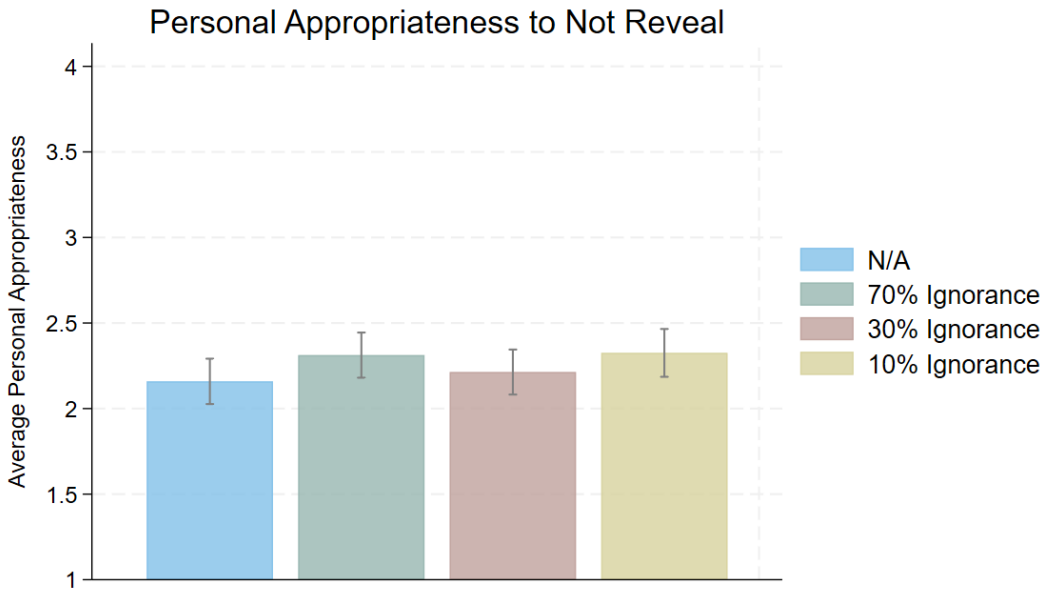
Hypothesis - Self-justifying Beliefs: Ignorant dictators will be more likely to rate a selfish ignorant action as appropriate than an informed dictator. Similarly, an informed dictator will be more likely to rate a prosocial reveal choice more favorably than an ignorant dictator.

The results indicate that personal appropriateness ratings remained statistically consistent across all treatment conditions, suggesting that individuals' moral evaluations did not depend on their information status. In contrast, perceptions of others' social appropriateness judgments varied systematically, indicating that dictators updated their beliefs about injunctive norms. Specifically, dictators in the *70% Ignorance* condition believed that others viewed choosing not to reveal information more favorably compared to other conditions, suggesting that perceived social norms influence second-order beliefs.

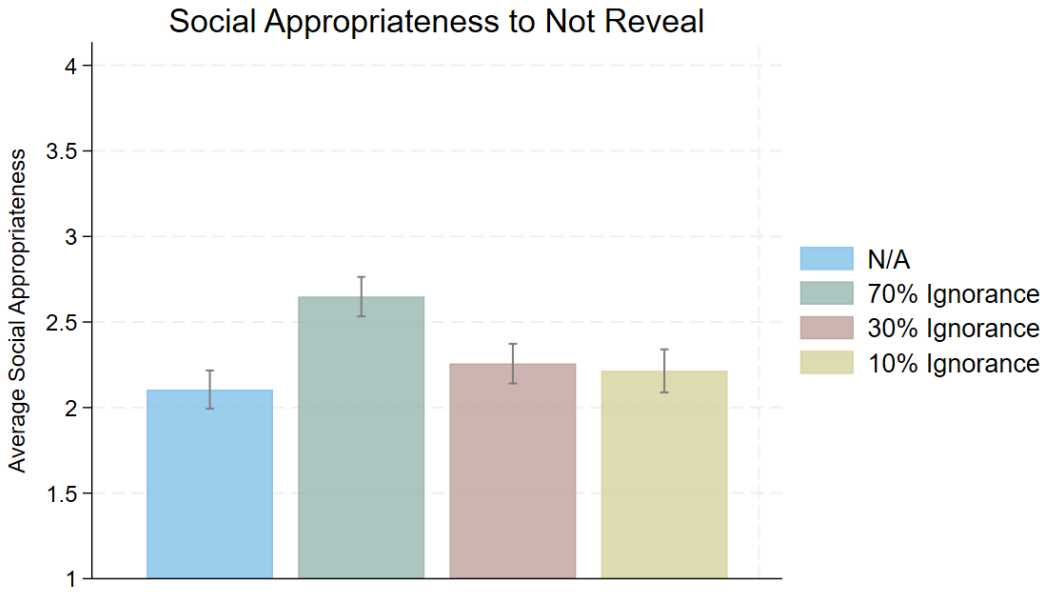
Figure 7 presents these results, highlighting the distinction between personal appropriateness judgments and incentivized social appropriateness perceptions.

J Regression-Based Balance Test

This appendix presents a regression-based balance test to examine the demographic characteristics of the sample. Most demographic covariates do not show statistically significant differences, except for student status and political affiliation. Some significance is also observed for certain religious and employment demographics, though these groups have relatively small sample sizes. Table 8 reports the regression results, which assess how subjects' demographic traits relate to their likelihood of avoiding information.



(a) Subject's personal belief in the appropriateness of choosing not to reveal.



(b) Subject's incentivized prediction of how they believe others rated the appropriateness of choosing not to reveal.

Figure 7: Perceived social appropriateness of information avoidance across norm treatments.

Notes: Bars indicate 95% confidence intervals.

Table 8: Demographics on Likelihood of Avoiding Information

	(1) avoided info	(2) avoided info	(3) avoided info
Age	-0.00361	-0.0107	0.00354
	(0.00455)	(0.00665)	(0.00667)
Female	0.0273	0.259*	-0.213
	(0.0928)	(0.138)	(0.136)
Black	0.127	-0.214	0.424**
	(0.153)	(0.243)	(0.211)
Hispanic	-0.223	-0.241	-0.230
	(0.198)	(0.314)	(0.272)
East Asian	-0.152	-0.209	-0.124
	(0.232)	(0.330)	(0.348)
Southeast Asian	0.191	-0.0386	0.323
	(0.261)	(0.383)	(0.384)
MENA	0.0139	0	0.453
	(0.594)	(.)	(0.697)
Other Race	-0.0571	-0.00424	-0.176
	(0.271)	(0.376)	(0.425)
less than $20,000	-0.00525	-0.191	0.230
	(0.271)	(0.393)	(0.397)
less than $40,000	-0.102	-0.374	0.161
	(0.287)	(0.418)	(0.419)
less than $60,000	-0.143	-0.604	0.318
	(0.288)	(0.424)	(0.416)
less than $80,000	-0.145	-0.635	0.267
	(0.295)	(0.440)	(0.423)
less than $100,000	-0.210	-0.661	0.272
	(0.316)	(0.474)	(0.447)
greater than $100,000	-0.169	-0.481	0.260
	(0.299)	(0.448)	(0.431)
Part-Time	-0.274*	-0.529**	-0.0517
	(0.143)	(0.209)	(0.214)
Self-Employed	-0.326**	-0.345	-0.468*
	(0.162)	(0.231)	(0.247)
Unemployed	-0.0316	-0.190	0.110
	(0.150)	(0.220)	(0.221)
Retired	-0.146	-0.296	0.0505
	(0.286)	(0.411)	(0.445)
Other Employment	-0.544**	-1.055***	-0.0284
	(0.213)	(0.327)	(0.307)
High school graduate	0.0752	0.843	-0.596
	(0.375)	(0.631)	(0.593)
Bachelors	0.189	0.999	-0.441
	(0.370)	(0.626)	(0.582)
Masters	0.215	1.044	-0.478
	(0.379)	(0.642)	(0.596)
Student	0.299*	0.0774	0.522**
	(0.153)	(0.234)	(0.219)
Independent	-0.196	-0.344*	-0.0951
	(0.125)	(0.183)	(0.181)
Democrat	-0.296**	-0.465**	-0.167
	(0.126)	(0.181)	(0.185)
Islam	-0.123	0	-0.362
	(0.471)	(.)	(0.567)
Judaism	0.375	1.170**	0.0182
	(0.307)	(0.530)	(0.427)
Hinduism	1.403**	1.576**	0
	(0.687)	(0.774)	(.)
Buddhism	-0.159	-0.643	0.373
	(0.443)	(0.641)	(0.719)
Agnosticism	0.0624	0.00138	0.211
	(0.126)	(0.181)	(0.187)
Atheism	-0.0591	0.0431	-0.0900
	(0.140)	(0.200)	(0.210)
Other Religion	0.356**	0.292	0.450**
	(0.149)	(0.223)	(0.213)
Condition	All	Known	Unknown
Observations	888	445	439

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

K Analysis of Demographic Covariates and Result on *Ignorance Expectant* Dictators

This appendix examines whether specific demographic variables systematically influence the likelihood of a participant being classified as an *reveal expectant* or *ignorance expectant* dictator. To assess this, we regress the probability of being a *ignorance expectant* dictator on a range of demographic covariates, including age, gender, income, education, political affiliation, and religious background.

The results indicate that no single demographic characteristic is disproportionately associated with *ignorance expectant* dictators. In other words, *ignorance expectant* and *reveal expectant* dictators appear to be relatively balanced across demographic groups, suggesting that Result 2 is not driven by an overrepresentation of any particular subgroup. This strengthens the conclusion that differences in norm responsiveness are behavioral rather than a function of underlying demographic composition.

Table 9 presents the full regression results.

L Appropriateness Valuation of Revealing and Not Revealing Between *Ignorance Expectants* and *Reveal Expectants*

Table 10 presents average ratings of personal and social appropriateness for revealing versus not revealing the recipient's payoff, disaggregated by belief type. Higher scores indicate greater perceived appropriateness. Personal appropriateness reflects each participant's own normative judgment (non-incentivized), while social appropriateness captures perceived injunctive norms and was incentivized using a coordination mechanism (i.e., matching the modal response of others). The results show that *reveal expectants* rate revealing as being more personally appropriate and not revealing as less appropriate compared to *ignorance expectants*. Interesting, both *ignorance expectants* and *reveal expectants* rated social appropriateness similarly, suggesting that *ignorance expectants* are aware of the social norm. These patterns suggest that individuals' normative views are aligned with their expectations about others' behavior, consistent with the idea that descriptive norms can shape both personal and social norm perceptions. Importantly, the distinction between incentivized social norms and non-incentivized personal norms echoes recent findings that personal norms play a central role in guiding economic be-

Table 9: Demographics on Being a *Ignorance Expectant* Versus *Reveal Expectant* Dictator

	(1)	(2)	(3)
	Ignorance Expectant	Ignorance Expectant	Ignorance Expectant
Ignorance Expectant			
Age	-0.00214	-0.0128	0.0104
	(0.00506)	(0.00790)	(0.00722)
Female	0.0155	0.0646	-0.0442
	(0.104)	(0.156)	(0.154)
Black	0.178	0.368	0.0806
	(0.165)	(0.247)	(0.236)
Hispanic	-0.105	0.229	-0.379
	(0.219)	(0.335)	(0.328)
East Asian	0.267	0.509	-0.0470
	(0.240)	(0.346)	(0.376)
Southeast Asian	0.188	0.500	-0.0550
	(0.286)	(0.390)	(0.485)
Other Race	-0.149	0.195	-0.438
	(0.316)	(0.416)	(0.591)
less than $20,000	0.336	0.324	0.265
	(0.315)	(0.499)	(0.431)
less than $40,000	0.272	0.566	-0.217
	(0.333)	(0.521)	(0.465)
less than $60,000	0.132	0.0528	0.111
	(0.335)	(0.532)	(0.456)
less than $80,000	-0.0887	0.183	-0.571
	(0.347)	(0.543)	(0.486)
less than $100,000	0.211	0.0329	0.130
	(0.363)	(0.584)	(0.490)
greater than $100,000	0.180	0.260	-0.0198
	(0.347)	(0.560)	(0.473)
Part-Time	-0.0881	-0.257	0.0102
	(0.157)	(0.232)	(0.239)
Self-Employed	-0.388**	-0.452*	-0.362
	(0.186)	(0.266)	(0.286)
Unemployed	-0.0583	-0.323	0.0878
	(0.167)	(0.245)	(0.246)
Retired	-0.00701	-0.196	0.0245
	(0.317)	(0.494)	(0.465)
Other Employment	-0.470*	-0.955**	-0.265
	(0.251)	(0.416)	(0.355)
High school graduate	-1.113	-0.435*	-0.102
	(0.940)	(0.250)	(0.693)
Bachelors	-0.980	-0.370	0.0315
	(0.941)	(0.232)	(0.684)
Masters	-0.801	0	0.182
	(0.948)	(.)	(0.697)
Student	-0.00992	-0.202	0.213
	(0.171)	(0.270)	(0.244)
Independent	-0.0797	-0.391*	0.194
	(0.140)	(0.219)	(0.203)
Democrat	-0.119	-0.156	-0.0254
	(0.140)	(0.206)	(0.208)
Judaism	-0.334	-0.218	-0.505
	(0.418)	(0.672)	(0.595)
Hinduism	-0.0831	-0.101	0
	(0.743)	(0.862)	(.)
Buddhism	-0.748	0	0.125
	(0.575)	(.)	(0.809)
Agnosticism	0.0796	0.169	-0.0360
	(0.141)	(0.207)	(0.210)
Atheism	0.205	0.403*	-0.00192
	(0.152)	(0.222)	(0.228)
Other Religion	0.238	0.663***	-0.175
	(0.165)	(0.249)	(0.245)
Condition	All	Known	Unknown
Observations	884	429	429

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

havior (Bašić and Verrina, 2024).

Table 10: Higher values indicate higher appropriateness ratings on a scale of 1 to 4. Social appropriateness was incentivized by matching the modal response of other participants, while personal appropriateness was non-incentivized.

	Ignorance Expectants	**Reveal Expectants**
Personal Appropriateness: Reveal	3.19	3.60
	(0.89)	(0.69)
Personal Appropriateness: Not Reveal	2.55	2.17
	(1.03)	(1.09)
Social Appropriateness: Reveal	3.31	3.36
	(0.79)	(0.77)
Social Appropriateness: Not Reveal	2.33	2.30
	(0.92)	(0.98)

M Exploratory Finding: Gender effects on Responses to Norms.

When looking at gender differences, women—but not men—appeared to resist normative forces in the *self-referential* frame *Known* treatment compared to the *socially* framed *Unknown* treatment. Figure 8 highlights this pattern. A two-sided difference-of-means test comparing the *Known* and *Unknown* conditions at the *10% Ignorance* level is significant at the 5% level ($t = 2.11, p < 0.05$). When pooling the *30% Ignorance* and *10% Ignorance* treatments, female dictators in the *Known* condition, where the frame is *self-centered*, were 14 percentage points less likely to reveal information ($t = 1.99, p < 0.05$). As a placebo test, this discrepancy does not appear for male dictators or female dictators in the *No Info* treatment, reinforcing the robustness of this unexpected result.

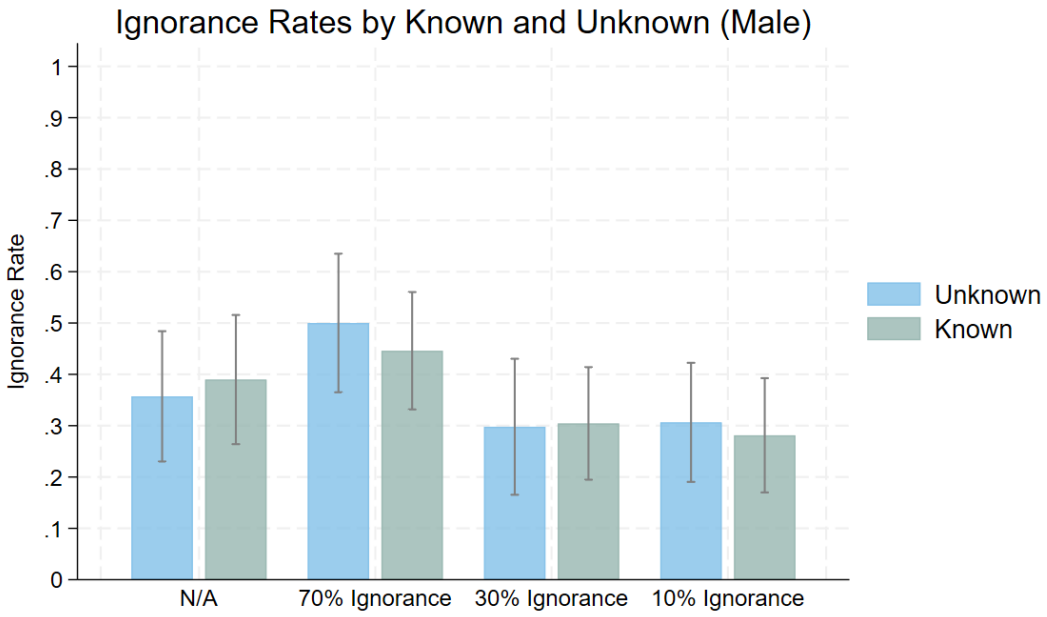
While previous work finds that women are more prosocial than men in dictator games (Eckel and P. J. Grossman, 1998), my results suggest that such gender differences are sensitive to the framing of social context. Specifically, when decisions are framed *self-referentially*, female participants exhibit a statistically significant resistance to normative influence—a reversal of the commonly assumed greater norm sensitivity among women (Croson and Gneezy, 2009). Rather than contradicting prior findings, these results indicate that prosocial behavior among women may be less about uncon-

ditional generosity and more about selective responsiveness to context, identity, and framing. This context dependence suggests that how agency and social expectations are framed plays a critical role. The resistance to normative influence under *self-referential* framing points to a potential cognitive or motivational difference in how women process norm-related cues, warranting further investigation into the underlying psychological mechanisms. Future research could explore whether this pattern persists beyond information avoidance behavior and across alternative social environments.

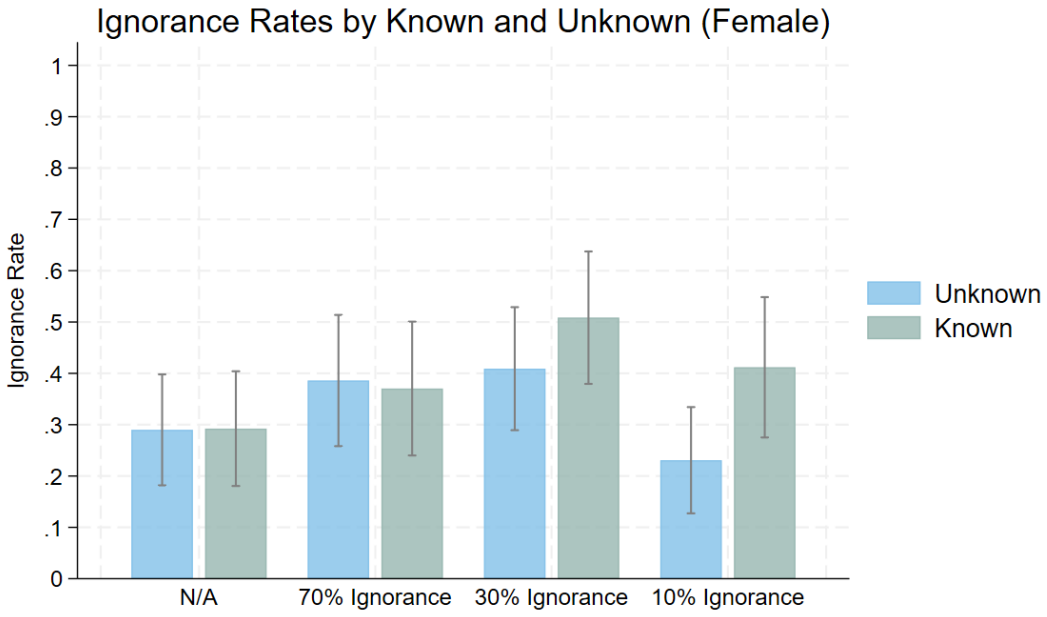
To further investigate this finding, I examined the choice process data on the interaction rate and time spent engaging with the mock-up interface between the *Known* and *Unknown* conditions. Figure 9 reports the choice process data. Under panel 9a, in the *Known* condition, women are 8 percentage points less likely than men to interact with the mock-up. However, among those who do engage, women are 14 percentage points less likely to reveal the payoff tables ($t = 2.52, p < 0.05$). This pattern does not emerge for men ($p = 0.88$), suggesting a gender-specific response to framing. Similarly, under panel 9b, dictators in the *Known* relative to the *Unknown* condition spent less time on the mock-up interface. Given the minimal differences in instructions between the *Known* and *Unknown* conditions, these differences in choice process data cannot be attributed to variations in text length.[17]

Examining subjects' moral universalism trust scores provides additional insight. In the *Known* condition, universalism negatively correlates with keeping the payoff tables hidden, whereas no such relationship emerges in the *Unknown* condition. This suggests that when individuals are explicitly aware of their decision-making role, those with stronger universalist tendencies are less likely to avoid information—potentially reflecting an increased sense of moral obligation toward out-groups. Similarly, those with lower universalist tendencies become more likely to behave selfishly. In contrast, when decision-making is *framed socially* (as in the *Unknown* condition), universalism appears less influential, implying that *self-referential* framing activates individuals' sensitivity to in-group versus out-group moral considerations. This aligns with Momsen and Ohndorf (2023), who find that political orientation affected the type of information avoidance, with some subjects avoiding learning about donation outcomes when self-serving motives are at play, while others avoid personal payoff information to maintain a moral commitment. Regression results and additional analysis reported in Appendix G.

[17] Appendix O contains screenshots of the experimental interface.



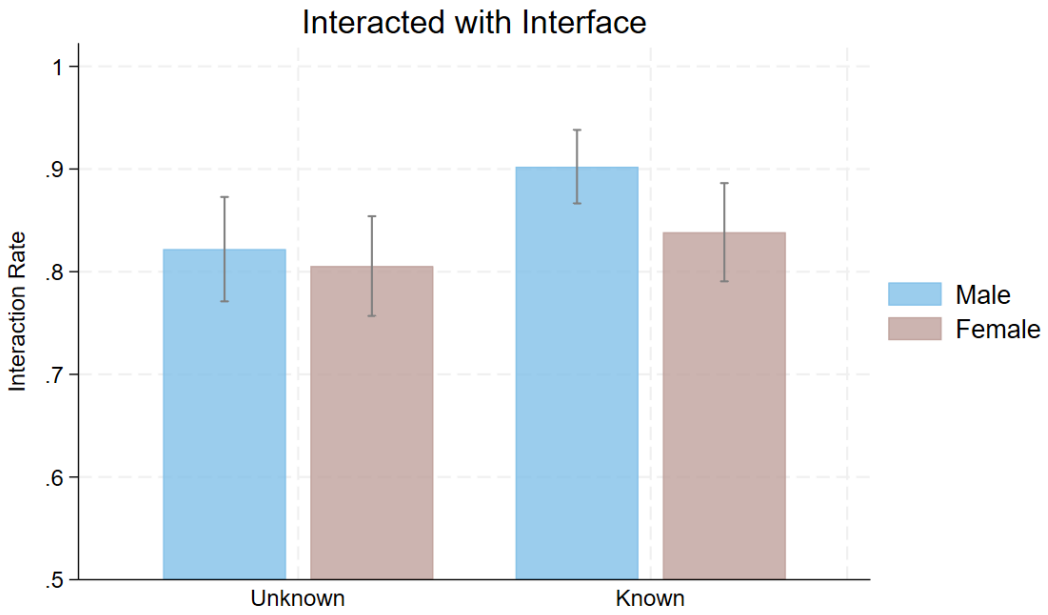
(a) Ignorance rates of male subjects.



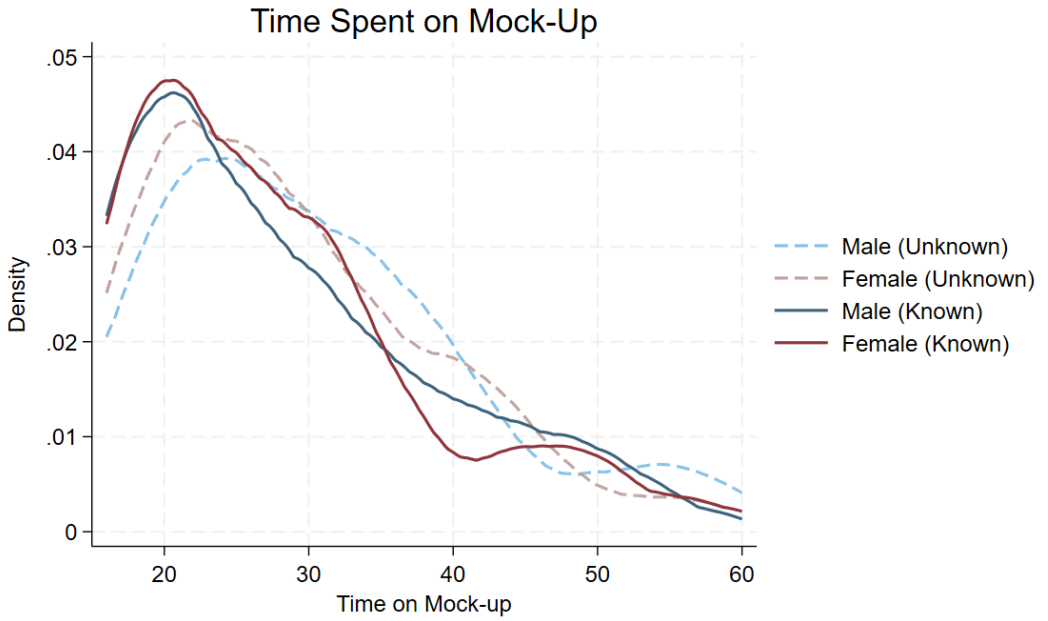
(b) Ignorance rates of female subjects.

Figure 8: Ignorance rates among male and female participants, split by treatment conditions.

Notes: Bars indicate 95% confidence intervals.



(a) Percentage of subjects interacting with the mock-up interface.



(b) Time spent on mock-up interface screen. X-axis truncated at 80 seconds for scale.

Figure 9: Choice process metrics (i.e., interface interaction, decision time) among male and female participants

Notes: Bars indicate 95% confidence intervals.

N Exploratory Finding: Effects of Belief Elicitation on Other's Behavior

As an additional exploratory analysis, I estimate the direct effect of belief elicitation by comparing this study's results with those of Z. Grossman, Hua, et al. (2025), which used the same Prolific subject pool and inclusion criteria.[18] In that study, dictators in treatment arms with nearly identical interfaces and instructions exhibited an average ignorance rate of 62%.[19] The only notable difference was the inclusion of a mock interface and belief elicitation. However, in the present study, when dictators were asked to predict others' behavior before making their own decision, ignorance dropped significantly from 62% to 32% in the *No Norm* condition. This sharp decline suggests that merely prompting individuals to consider others' choices reduces strategic ignorance, likely by reinforcing the social dimension of decision-making.

To ensure this effect is not merely due to the presence of a mock interface, I also compare these results with those of Mol, Soraperra, and van der Weele (2025), who implemented a test round in their version of the moral wiggle-room game also using a Prolific subject pool and observed a 60% ignorance rate.[20]

The comparison highlights the significant impact of belief elicitation on reducing strategic ignorance, suggesting that prompting individuals to consider others' behavior can promote information acquisition, possibly by increasing the salience of the decision, as in Z. Grossman (2014), who framed ignorance as an active choice, and Z. Grossman, Hua, et al. (2025), who required dictators to explicitly decide whether to acquire information before proceeding. This finding indicates that interventions designed to curb strategic ignorance may be effective even in the absence of direct norm enforcement, as increasing awareness of collective behavior alone can shape decision-making. Moreover, asking individuals to predict others' choices may subtly reinforce the injunctive norm that links information-seeking with prosocial behavior. By making this norm more salient, belief elicitation may help counteract potential boomerang effects, which could otherwise lead dictators to embrace ignorance as a justification for self-serving decisions (Schultz et al., 2007).

[18] This effect was not hypothesized in the pre-registration, as the magnitude of the shift in ignorance rates was unexpected, and it was unclear whether belief elicitation would meaningfully influence behavior.

[19] Appendix P provides the interface details from Z. Grossman, Hua, et al. (2025).

[20] Recipient types varied: Z. Grossman, Hua, et al. (2025) matched dictators with another Prolific user, while Mol, Soraperra, and van der Weele (2025) used a charity recipient.

These results align with a mental model framework, wherein individuals rely on internalized cognitive structures to interpret decision environments and guide choices. Mental models shape both information processing and responses to framing effects, acting as cognitive filters that influence attention allocation and the weight given to moral norms. In addition to being asked to formulate beliefs on the behavior of others, when decisions are framed *socially* (*Unknown* condition), individuals activate a norm-driven mental model, increasing deliberation and conformity to perceived norms. Conversely, when framed as a *self-referential*, personal choice (*Known* condition), individuals adopt a self-focused mental model that streamlines decision-making and promotes information avoidance by deprioritizing social factors. This perspective parallels Charness and Rabin (2002), who show that people prioritize social welfare when fairness concerns are salient but withdraw moral consideration when others act selfishly—a mechanism that may underlie the observed differences in information acquisition. It also helps explain van der Weele et al. (2014), who found reciprocal behavior persisted despite moral wiggle-room. Future research should explore whether these framing effects arise from cognitive processing styles, socialization, or the flexibility of mental model activation.

O Experiment Interface

This appendix contains screenshots of the experimental interface. When appropriate, alternative versions of the same screen corresponding to different treatment arms are provided. For the *Known* condition, when the subject is aware that they will later participate as the dictator, the interface refers to the dictator as "you." In the *Unknown* condition, before making the allocation decision, the interface refers to the dictator as "Person 1."

Page 1 Instructions

All Treatments

Thank you for participating in this study. Depending on the decisions you make, you can earn additional bonus payments. All payment amounts will be stated in terms of Experimental Currency Units (ECUs) converted at a **rate of 2 ECUs = $1**.

We will provide you with 5 ECUs ($2.50) for the completion of the study. Additional bonus ECUs earned will be provided within 72 hours after the study has been completed. You may withdraw from the study at any time. However, if you end the study before completion, you will forfeit all payment, including any bonuses earned.

Please do not leave your computer, and keep your attention on this browser tab until the study has been completed. There will be timers and attention checks. If you remain idle for an extended amount of time, your participation in the study will be terminated.

Continue

Page 2 Instructions

Unknown treatment

In studies like this one, we typically present people with a scenario in which one person chooses between two options that determine their own payoff and that of another person. We are going to describe one of these scenarios to you.

Afterwards, you will make some predictions about what you think other people have done in that scenario.

Here is an example to help you understand what is going on. After you answer some questions about it to confirm your understanding, we will describe the actual scenario to you.

There are two people: Person 1 and Person 2. Person 1 picked between options A and B to decide how much they and a second individual, Person 2, will receive.

Person 2 does not make any decisions.

In this example, the payoff table below shows that if Person 1 chooses B, then they get 3 ECUs and Person 2 receives 4 ECUs.

Example Payoff Table

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	1	2
B	3	4

Question 1: In this example, if Person 1 chooses A, then how much does Person 1 receive?

Question 2: In this example, if Person 1 chooses A, then how much does Person 2 receive?

button appears in 13 seconds

Page 2 Instructions

Known treatment

In studies like this one, we typically present people with a scenario in which one person chooses between two options that determine their own payoff and that of another person. We are going to describe one of these scenarios to you.

Afterwards, you will make some predictions about what you think other people have done in that scenario. Then you will face that scenario yourself and make your own decisions.

Here is an example to help you understand what is going on. After you answer some questions about it to confirm your understanding, we will describe the actual scenario to you.

There are two people: Person 1 and Person 2. Person 1 picked between options A and B to decide how much they and a second individual, Person 2, will receive.

Person 2 does not make any decisions.

In this example, the payoff table below shows that if Person 1 chooses B, then they get 3 ECUs and Person 2 receives 4 ECUs.

Example Payoff Table

You Choose	You Receive	Person 2 Receives
A	1	2
B	3	4

Question 1: In this example, if you as Person 1 choose A, then what payment do you receive?

Question 2: In this example, if you as Person 1 choose A, then what payment does Person 2 receive?

button appears in 13 seconds

Page 3 Instructions

Unknown treatment

Now we are going to describe to you the actual scenario that other participants faced. At the bottom of this page, you will answer some questions to confirm your understanding.

In the actual scenario, Person 1 will see the Hidden Payoff Table which hides Person 2's payoffs from Person 1. However, Person 1 may **optionally** reveal the Hidden Payoff Table to determine if it is one of **two possible** versions of the payoff tables (shown below): Payoff Table 1 or Payoff Table 2, with an **equal chance** of it being either.

The computer will choose the actual payoff table ahead of time. Revealing the payoff table only shows which version is being used. **Person 2 will not know if Person 1 chooses to reveal** the Hidden Payoff Table.

According to these two payoff tables, Person 1 will always receive 6 ECUs if they pick option A and 5 ECUs if they pick option B. Person 2 will receive either 1 or 5 ECUs depending on the actual payoff table.

Hidden Payoff Table

Person 1 Choose	Person 1 Receive	Person 2 Receives
A	6	???
B	5	???

Payoff Table 1

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	6	1
B	5	5

Payoff Table 2

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	6	5
B	5	1

Several people in the other study acting as Person 1 made a decision whether or not to reveal the Hidden Payoff Table. We will ask you to **predict the percentage of those people acting as Person 1 who chose to reveal the table.**

Question 3 : If Person 1 did not reveal the Hidden Payoffs Table and chooses B, what will Person 2 receive?

- Person 2 will always get 6
- Person 2 will always get 5
- Person 2 will always get 1
- Person 2 will get either 5 or 1
- Person 2 will get either 6 or 5

Question 4: If Person 1 chooses A, what will Person 1 receive?

- Person 1 will always get 6
- Person 1 will always get 5
- Person 1 will always get 1
- Person 1 will get either 5 or 1
- Person 1 will get either 6 or 5

button appears in 14 seconds

Page 3 Instructions

Known treatment

Now we are going to describe to you the actual scenario that other participants faced and that you will later participate in. At the bottom of this page, you will answer some questions to confirm your understanding.

In the actual scenario, you will see the Hidden Payoff Table which hides Person 2's payoffs from you. However, you may **optionally** reveal the Hidden Payoff Table to determine if it is one of **two possible** versions of the payoff tables (shown below): Payoff Table 1 or Payoff Table 2, with an **equal chance** of it being either.

The computer will choose the actual payoff table ahead of time. Revealing the payoff table only shows which version is being used. **Person 2 will not know if you choose to reveal** the Hidden Payoff Table.

According to these two payoff tables, you will always receive 6 ECUs if you pick option A and 5 ECUs if you pick option B. Person 2 will receive either 1 or 5 ECUs depending on the actual payoff table.

Hidden Payoff Table

You Choose	You Receive	Person 2 Receives
A	6	???
B	5	???

Payoff Table 1

You Choose	You Receive	Person 2 Receives
A	6	1
B	5	5

Payoff Table 2

You Choose	You Receive	Person 2 Receives
A	6	5
B	5	1

Several people in the other study acting as Person 1 made a decision whether or not to reveal the Hidden Payoff Table. We will ask you to **predict the percentage of those people acting as Person 1 who chose to reveal the table.** After your prediction, you will act as Person 1 in a similar experiment.

Question 3: If you did not reveal the Hidden Payoffs Table and choose A, what will Person 2 receive?

- Person 2 will always get 6
- Person 2 will always get 5
- Person 2 will always get 1
- Person 2 will get either 5 or 1
- Person 2 will get either 6 or 5

Question 4: If you choose A, what will you receive?

- You will always get 6
- You will always get 5
- You will always get 1
- You will get either 5 or 1
- You will get either 6 or 5

button appears in 13 seconds

Page 4 Instructions

Unknown treatment

We want to give you a chance to familiarize yourself with the choice interface as it was presented to the other participants to help you understand it before you make your prediction.

Person 2's payoffs are hidden by "???" but Person 1 can choose to reveal the payoffs before choosing between option A and option B. **Choosing to reveal is free.**

Please take a moment to imagine what Person 1 would do in this scenario. The interface below is just a mock-up. The buttons do not do anything if clicked.

Hidden Payoff Table

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	6	???
B	5	???

- A
- B
- Reveal Payoff Table

button appears in 8 seconds

Page 4 Instructions

Known treatment

We want to give you a chance to familiarize yourself with the choice interface as it was presented to the other participants to help you understand it before you make your prediction.

Person 2's payoffs are hidden by "???" but Person 1 can choose to reveal the payoffs before choosing between option A and option B. **Choosing to reveal is free.**

Please take a moment to imagine what you would do in this scenario. The interface below is just a mock-up. The buttons do not do anything if clicked.

Hidden Payoff Table

You Choose	You Receive	Person 2 Receives
A	6	???
B	5	???

A

B

Reveal Payoff Table

button appears in 7 seconds

Page 5 Instructions

Unknown treatment

Hidden Payoff Table

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	6	???
B	5	???

This is the Hidden Payoffs Table that participants acting in the role of Person 1 sees when making their decision. Person 2's payoffs are hidden by "???" but **Person 1 can choose to reveal the payoffs before choosing between option A and option B. Choosing to reveal is free.**

What percentage of people acting in the role of Person 1 do you believe revealed the payoff tables?

If your prediction is more accurate than the majority of the other participants, you will earn an additional ECU.

Use the slider and plus or minus buttons below to make your prediction.

No one revealed — Everyone revealed

\-

Percentage of people who **revealed**: %

\+

button appears in 13 seconds

Page 5 Instructions

Known treatment

Hidden Payoff Table

Person 1 Chooses	Person 1 Receives	Person 2 Receives
A	6	???
B	5	???

This is the Hidden Payoffs Table that participants acting in the role of Person 1 sees when making their decision. Person 2's payoffs are hidden by "???" but **Person 1 can choose to reveal the payoffs before choosing between option A and option B. Choosing to reveal is free.**

Before you participate as Person 1, what percentage of people acting in the role of Person 1 do you believe revealed the payoff tables?

If your prediction is more accurate than the majority of the other participants, you will earn an additional ECU.

Use the slider and plus or minus buttons below to make your prediction.

button appears in 18 seconds

Page 6 Instructions

Unknown treatment

Now we will ask you to act in the role of Person 1. Another Prolific participant will be matched and assigned the role of Person 2. You will earn additional ECUs based on your decision. Recall that these are the two possible payoff tables, **with both being equally likely.**

Payoff Table 1

You Choose	You Receive	Person 2 Receives
A	6	1
B	5	5

Payoff Table 2

You Choose	You Receive	Person 2 Receives
A	6	5
B	5	1

You will not know which payoff table is being used, but you can choose to reveal which payoff table is being used. We will ask you to make your decision on the next page.

Continue

Page 6 Instructions

Known treatment

Now it is time for you to make your decision as Person 1. Another Prolific participant will be matched and assigned the role of Person 2. You will earn additional ECUs based on your decision. Recall that these are the two possible payoff tables, **with both being equally likely.**

Payoff Table 1

You Choose	You Receive	Person 2 Receives
A	6	1
B	5	5

Payoff Table 2

You Choose	You Receive	Person 2 Receives
A	6	5
B	5	1

You will not know which payoff table is being used, but you can choose to reveal which payoff table is being used. We will ask you to make your decision on the next page.

Continue

Page 7 Instructions

30% Ignorance treatment - Overestimated Ignorance

A majority: 70% or 7 out of 10 people chose to reveal the payoff tables.

You predicted that **53% of people revealed** the payoff tables.

You **underestimated** the amount of people who had revealed the payoff tables. People were less likely to keep Person 2's payoffs hidden than you had predicted.

You are Person 1. If you choose "Reveal Payoff Table", the next page will reveal the payoffs Person 2 will receive when you choose between A and B.

You Choose	You Receive	Person 2 Receives
A	6	???
B	5	???

A

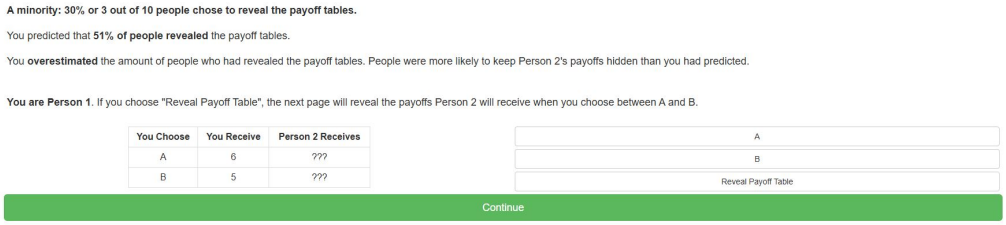
B

Reveal Payoff Table

Continue

Page 7 Instructions

70% Ignorance treatment - Underestimated Ignorance



Page 7 Instructions

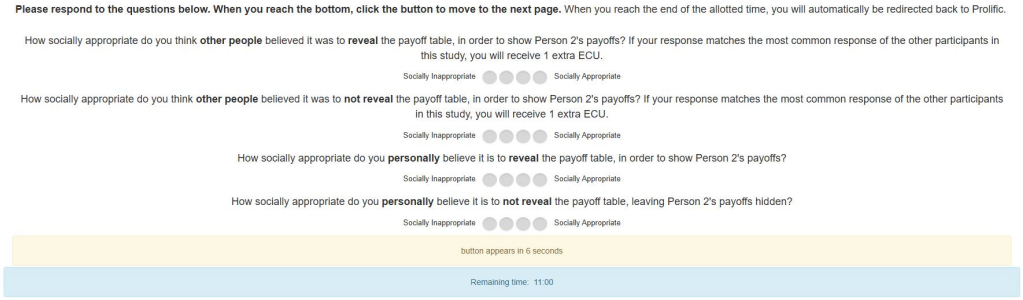
No Norm treatment

Page 7 Instructions

Revealed payoffs

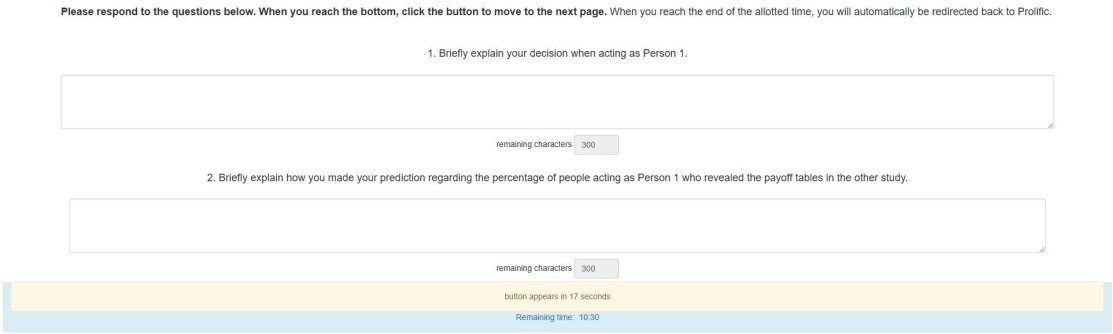
Page 8 Instructions

Appropriateness Rating



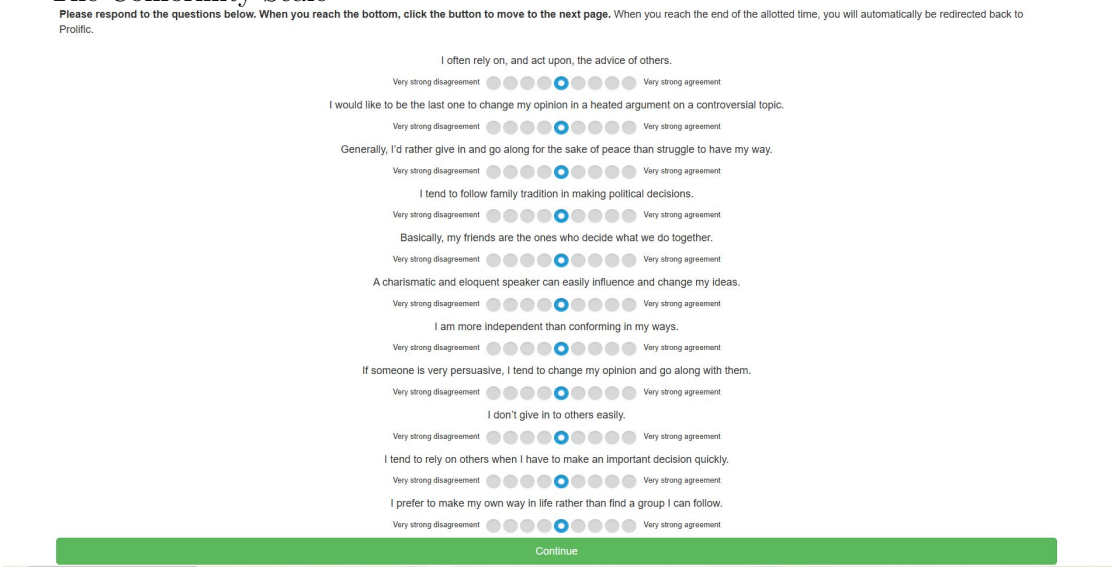
Page 9 Instructions

Free Response Explanation



Page 10 Instructions

The Conformity Scale



Page 11 Instructions

Demographics Survey

Please respond to the questions below. When you reach the bottom, click the button to move to the next page. When you reach the end of the allotted time, you will automatically be redirected back to Prolific.

1. What is your assigned sex at birth?

- Male
- Female
- Other

2. How many years old are you?

3. What is your race, ethnic, or ancestral background?

- White
- Black or African-American
- Hispanic or Latino/Latina/Latinx
- East Asian
- Southeast Asian
- Middle Eastern or North African
- Other

4. What is your annual income before taxes?

- $0
- $20,000 or less
- $40,000 or less
- $60,000 or less
- $80,000 or less
- $100,000 or less
- Greater than $100,000

5. What is your employment status?

- Full-Time
- Part-Time
- Self-Employed
- Unemployed
- Retired
- Other

6. What is your highest educational attainment?

- Less than high school
- High school graduate
- Bachelor's degree
- Master's or professional degree
- Doctoral degree

7. Are you currently a student?

- Yes
- No

8. In terms of your political leanings, how would you rate yourself?

Extremely liberal ○ ○ ○ ○ ○ ○ ○ Extremely conservative

9. With what political party, if any, do you sympathize?

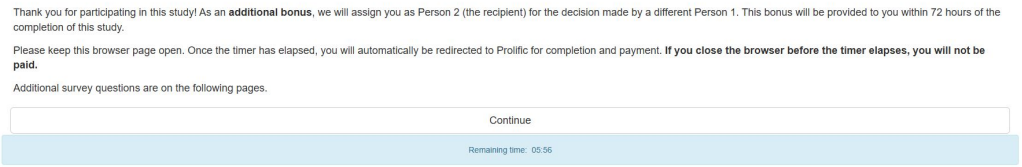
- Democrat
- Independent/none
- Republican

10. With which religious affiliation do you most identify with?

- Christianity
- Islam
- Judaism
- Hinduism
- Buddhism
- Agnosticism
- Atheism
- Other

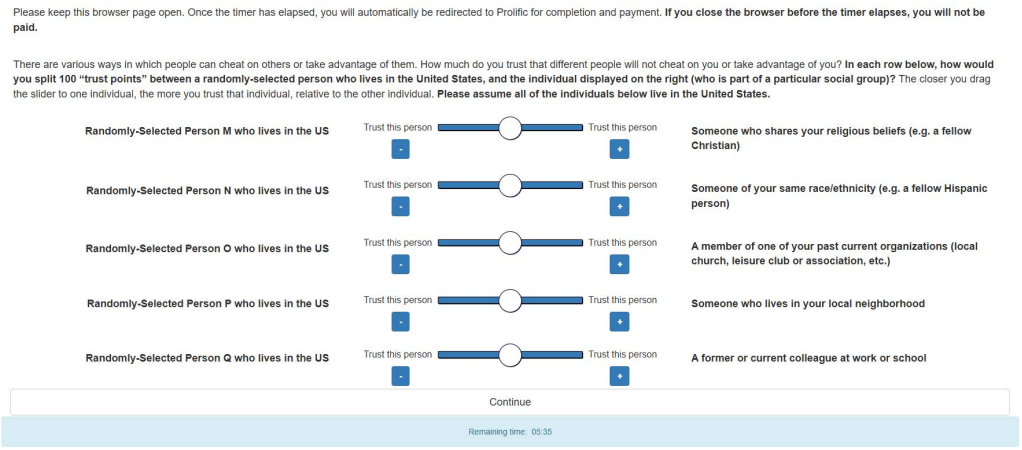
Page 12 Instructions

End Screen



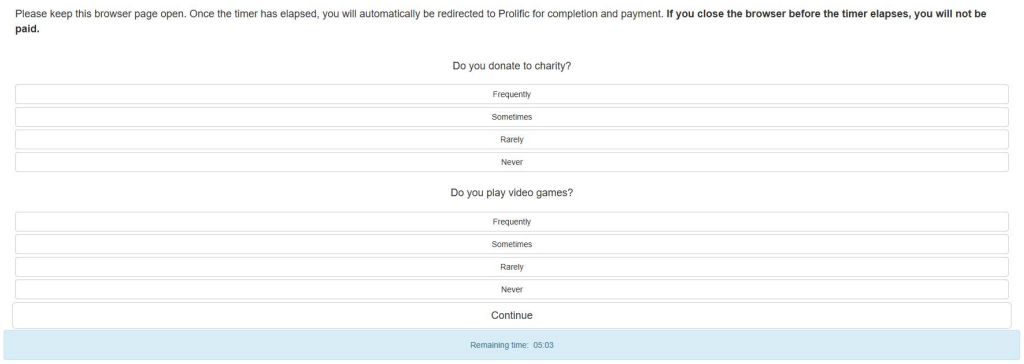
Page 13 Instructions

Optional Moral Universalism Trust Scale



Page 14 Instructions

Optional Questions

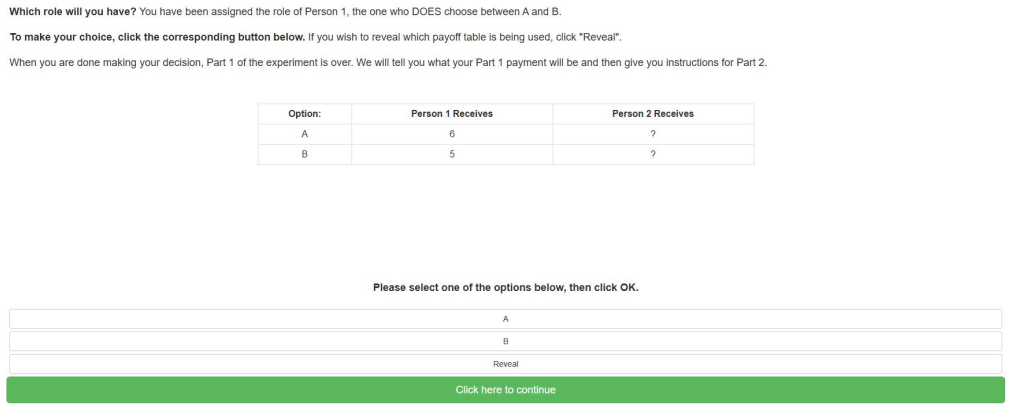


P Interface From Other Study

The screenshots show the interface from Z. Grossman, Hua, et al. (2025) across two different treatment arms. The first interface is the standard moral wiggle-room interface while the second interface involved subjects making an information decision before the option to make an allocation decision was available. The interface from this study is most comparable to the first interface. Thus, Result 1 specifically compares the first interface with the *Known - No Norms* and *Unknown - No Norms* conditions.

Program Interface from Other Experiment

Interface from Z. Grossman, Hua, et al., 2025



Which role will you have? You have been assigned the role of Person 1, the one who DOES choose between A and B.

To make your choice, click the corresponding button below. If you wish to reveal which payoff table is being used, click "Reveal".

When you are done making your decision, Part 1 of the experiment is over. We will tell you what your Part 1 payment will be and then give you instructions for Part 2.

Option:	Person 1 Receives	Person 2 Receives
A	6	?
B	5	?

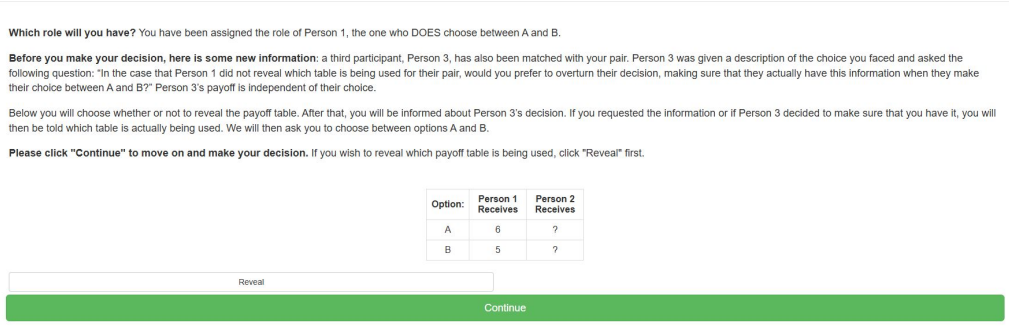
Please select one of the options below, then click OK.

A

B

Reveal

Click here to continue



Which role will you have? You have been assigned the role of Person 1, the one who DOES choose between A and B.

Before you make your decision, here is some new information: a third participant, Person 3, has also been matched with your pair. Person 3 was given a description of the choice you faced and asked the following question: "In the case that Person 1 did not reveal which table is being used for their pair, would you prefer to overturn their decision, making sure that they actually have this information when they make their choice between A and B?" Person 3's payoff is independent of their choice.

Below you will choose whether or not to reveal the payoff table. After that, you will be informed about Person 3's decision. If you requested the information or if Person 3 decided to make sure that you have it, you will then be told which table is actually being used. We will then ask you to choose between options A and B.

Please click "Continue" to move on and make your decision. If you wish to reveal which payoff table is being used, click "Reveal" first.

Option:	Person 1 Receives	Person 2 Receives
A	6	?
B	5	?

Reveal

Continue