MPRA
Munich Personal RePEc Archive

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

Hua, Tony

University of California, Merced

1 April 2025

Online at https://mpra.ub.uni-muenchen.de/125115/
MPRA Paper No. 125115, posted 03 Jul 2025 14:02 UTC

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

Tony Hua*[1]

[1]University of California, Merced

June 2025

Abstract

People often avoid information as a way to justify selfish behavior through plausible deniability. 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 acquire information when exposed to norms favoring transparency versus norms favoring ignorance. Such norm conforming behavior differed between *ignorance expectants*–those who expected ignorance–versus *reveal expectants*–those who expected transparency. Furthermore, there is no evidence that participants formed or adjusted their beliefs in a self-serving manner to justify selfish choices, either ex-ante or ex-post. These findings suggest that strategic ignorance responds to descriptive norms, but not through motivated belief distortion.

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

JEL Codes: C72, C91, D8, D9

*Email: thua5@ucmerced.edu. 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, Richard Patterson, and two anonymous reviewers 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 form of willful ignorance—known as moral wiggle-room—allows people to maintain a positive self-image while acting selfishly Dana, Weber, and Kuang, 2007; Feiler, 2014; Z. Grossman, 2014; Z. Grossman and Van der Weele, 2017; Exley and Kessler, 2023. In organizational and institutional settings, actors often eschew morally relevant information to preserve plausible deniability, particularly when transparency would entail reputational or ethical costs. 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. When considering whether to reveal information, a person 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 transparency, depending on what individuals expect others to do. In this way, beliefs about social norms can shape whether or not ignorance is a socially acceptable excuse.

This paper investigates how individuals' beliefs about the prevalence of information avoidance—i.e., descriptive norms—influence their own tendency to remain strategically ignorant. Social norms are known to shape economic behavior in diverse settings, including tax compliance, market participation, and prosocial conduct (Tomasello, Kruger, and Ratner, 1993; Akerlof, 1980). A central distinction lies between descriptive norms, which capture what people believe others typically do, and injunctive norms, which indicate what is socially approved (Bicchieri, Dimant, Gelfand, et al., 2023). While prior work highlights how norms guide behavior, less is known about whether individuals adopt self-serving beliefs about these norms to rationalize ignorance. For example, individuals

might excuse their behavior by assuming that ignorance is widespread, or reinterpret normative expectations in ways that favor inaction. Such norm-based justifications may play a critical role in sustaining ethical blind spots in economic decision-making.

A body of work suggests that individuals process information, often subconsciously, in ways that reinforce their pre-existing perspectives and social identities (Shepperd, Malone, and Sweeny, 2008). Self-serving biases distort fairness perceptions in negotiations (Babcock and Loewenstein, 1997), motivated reasoning enables individuals to justify self-interested behavior while maintaining a moral self-image (Rustichini and Villeval, 2014; Gino, Norton, and Weber, 2016), and selective belief distortion allows individuals to rationalize their actions without explicitly rejecting moral standards (Bicchieri, Dimant, and Sonderegger, 2023). These tendencies suggest that the willingness to avoid information may not simply reflect a desire to avoid moral responsibility, but also the extent to which individuals perceive strategic ignorance as socially acceptable. 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. This study tests whether such motivated cognition occurs.

This study contributes to the literature on strategic ignorance by examining how observed descriptive norms influence individuals' willingness to avoid morally relevant information. Building on Spiekermann and Weiss (2016), who model norm compliance as belief-contingent, I instead test whether perceived descriptive norms shape ignorance through social coordination or justification. 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 receive one of several norm messages, either stating that 10%, 30%, or 70% of others avoided information, or no feedback at all. This design allows me to cleanly identify how expectations—both self-generated and externally provided—influence strategic ignorance. Section 2 further describes the experimental procedures.

By manipulating information about social norms surrounding information avoidance, this experiment 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? The findings show that both pre-existing beliefs about norms and experimentally provided norm feedback influence strategic ignorance. Participants were more likely to seek information when exposed to norms favoring transparency, with *ignorance expectants*—those who initially believed avoidance was common—showing increased responsiveness, but only in the 10% ignorance environment. Furthermore, there is no evidence that participants distorted their beliefs in a self-serving way, either before or after making their own decisions.

While some subgroup-level shifts indicate sensitivity to norm feedback, there is little evidence of broad convergence or "pooling" towards either universal ignorance or transparency norms. Instead, norm effects appear concentrated among specific belief types. *Ignorance expectants* adjusted their behavior when feedback strongly conflicted with their expectations, whereas *reveal expectants*–those who anticipated widespread transparency–were limited in their responsive to norms and demonstrated a strong preference for information. This suggests that norm-based convergence toward ignorance is more likely within homogenous communities where expectations are aligned.

These findings contribute to the understanding of how social norms shape economic

decision-making. By demonstrating that strategic ignorance is sensitive to beliefs about others' behavior, this paper provides insight into the conditions under which individuals engage in information avoidance and the potential policy interventions that could mitigate it. More broadly, these results inform debates on moral decision-making, belief formation, and norm-based interventions in economic environments. Detailed findings reported in Section 3.

2 Experiment Design

This section details the experimental design, key treatment manipulations, and hypotheses. 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 for themselves and a recipient. 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.

The experiment introduces two modifications to the standard moral wiggle-room paradigm. First, it elicits participants' beliefs about others' information avoidance before they make their own decision. This temporal ordering ensures that beliefs are not influenced by subjects' own behavior, allowing for cleaner identification of how perceived descriptive norms shape subsequent choices. Second, it manipulates incentives to form self-serving beliefs. 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. 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. This framing variation is designed to test whether making one's decision-making role salient from the outset alters the way beliefs are formed and how norms influence behavior.

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

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.

2.1 Mock-Up Allocation and Belief Elicitation

In the first section of the experiment, subjects are introduced to the moral wiggle-room game (Dana, Weber, and Kuang, 2007) and complete comprehension tests. They then preview the allocation interface before predicting the percentage of dictators who revealed payoff tables from a previous and similar experiment. Specifically, subjects answer: "What percentage of people acting as Person 1 (the dictator) do you believe

revealed the payoff tables?"[1] 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.

2.2 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.[2] They are explicitly informed that recipients will not know whether the dictator chose to reveal the information. Depending on the treatment condition, dictators then receive feedback on their earlier prediction: they are shown their estimate, whether they overestimated or underestimated the actual reveal rate, and are provided with one of three descriptive norm cues: *10% Ignorance*, *30% Ignorance*, or *70% Ignorance*, reflecting how often others in a similar study chose to remain ignorant. Full details on these norm treatments are provided in Section 2.4. In the *No Norm* control condition, participants similarly make their prediction but receive no norm cues when making their allocation decision.

2.3 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

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

[2] 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.

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.

2.4 Manipulations

Using a 2×4 between-subjects design, the experiment investigates how individuals form their beliefs and how they behave in response to the behavior of others. In the *Unknown* condition, participants are introduced to the task by through the description of a generic actor, "Person 1," who decide whether or not to reveal the payoff information when making choosing between their payoff options. By contrast, in the *Known* condition, participants are told from the outset that they will act as the decision-maker and are asked to predict what others would do, potentially triggering self-serving beliefs.

Norm information was varied across four conditions: *No Norm* (no information), *70% Ignorance* (3 of 10 acquired information), *30% Ignorance* (7 of 10 acquired information), and *10% Ignorance* (9 of 10 acquired information). In the norm feedback

conditions, participants were shown a message indicating the observed rate of information acquisition, phrased as: "A majority (minority): X% or N out of 10 people chose to reveal the payoff tables," with the majority/minority label reflecting whether the observed rate was above or below 50%, X indicating the percentage, and N as the ratio.

These manipulations test whether individuals adjust their beliefs when anticipating justification needs and whether their decisions respond to perceived descriptive norms. Figure 1 summarizes the experimental procedures.

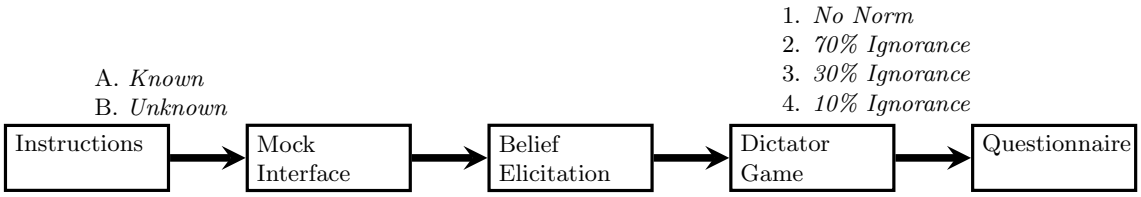


Figure 1: Experimental Procedures

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 and recipient type resulted in a notably low rate of ignorance.

Subjects were asked to predict behavior from prior experiments without being shown the exact interface or full experimental details. This simplification is justified because the study focuses on participants' perception of social norms, not precise recall or replication

of past behavior. Participants were told that they were predicting behavior in a similar but not identical decision environment. The *30%* and *70% Ignorance* conditions were selected to represent clear directional norms—where a strong majority either acquires or avoids information. In addition, since pilot participants typically estimated ignorance rates near 20%, the *10% Ignorance* condition was included to test how participants react when they overestimated ignorance.[3]

2.5 Hypotheses

This study examines how beliefs and social norms influence strategic ignorance by testing four hypotheses. These hypotheses explore the extent to which individuals adjust their information avoidance behavior in response to norms, their prior beliefs, and self-serving justifications. Collectively, these hypotheses evaluate the more general notion that multiple ignorance norms can exist due to a combination of inconsistent beliefs and responsiveness to norm environments. All hypotheses were pre-registered on 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 of transparency. 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

[3]See Appendix A.2 for further details.

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* (who expected low ignorance) and *ignorance expectants* (who expected high ignorance) 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.

Individuals may form self-serving beliefs about social norms in anticipation of their own behavior. If someone expects to benefit from avoiding information, they may inflate their belief about how common ignorance is to justify that choice in advance. To test this, I compare elicited beliefs between participants who know they will be making the allocation decision (the *Known* condition) and those who do not yet know their 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 report higher ignorance rates for social norms.

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[4], 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

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

participants received a base payment of \$2.50–\$2.70 for completing the study.[5] 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.[6] 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.1.

Table 2: Abridged subject demographics breakdown

	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)

[5] 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.

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

3.2 Primary Analysis

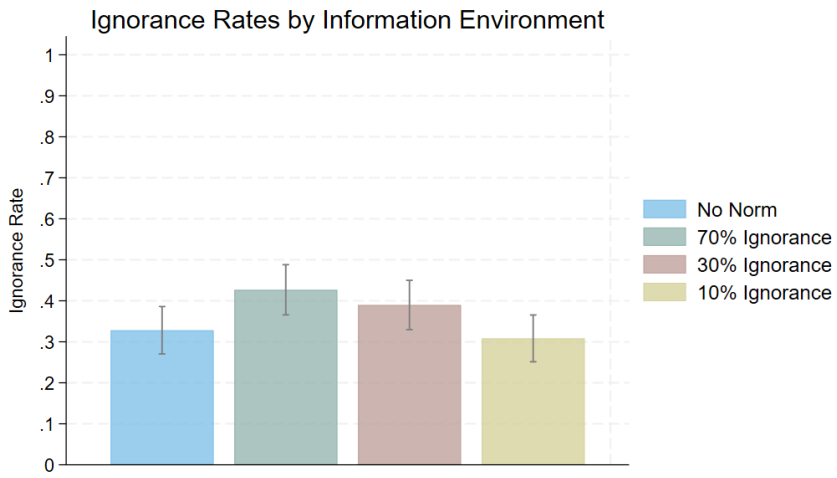
To confirm that subjects behaved in line with previous studies, I examined dictators' allocation decisions. Among the 370 who remained ignorant, 85% chose the self-serving option (6 over 5), consistent with past evidence that ignorance facilitates selfish behavior. Informed dictators in the aligned-interest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

Figure 2: Dictator Ignorance Rates Across Treatments



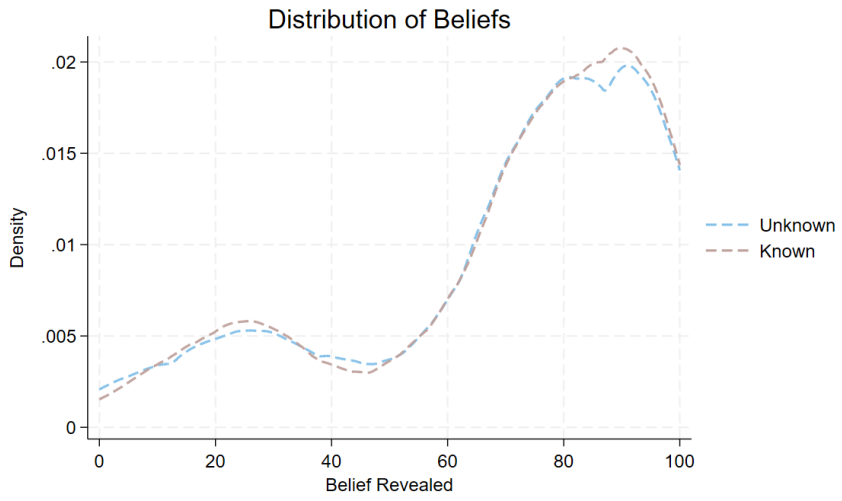
Notes: Bars indicate 95% confidence intervals.

remain uninformed. These terms reflect participants' beliefs about the prevailing descriptive norm—whether the majority seeks or avoids information[7]. 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.

The difference in baseline behavior between the two groups is stark. Under the *No Norm* condition, *ignorance expectant* dictators were 40 percentage points more likely to remain ignorant than *reveal expectants*. Figure 4 illustrates this gap by showing ignorance rates across treatment arms for both belief types. Appendix A.8 explores whether demographic differences account for this pattern, while Appendix A.9 examines differences in how each group rated the appropriateness of remaining ignorant.

[7]Appendix A.8 shows a demographics regression indicating no notable difference in the background of *ignorance expectants*.

Figure 3: Distribution of Subjects' Beliefs on Reveal Rates

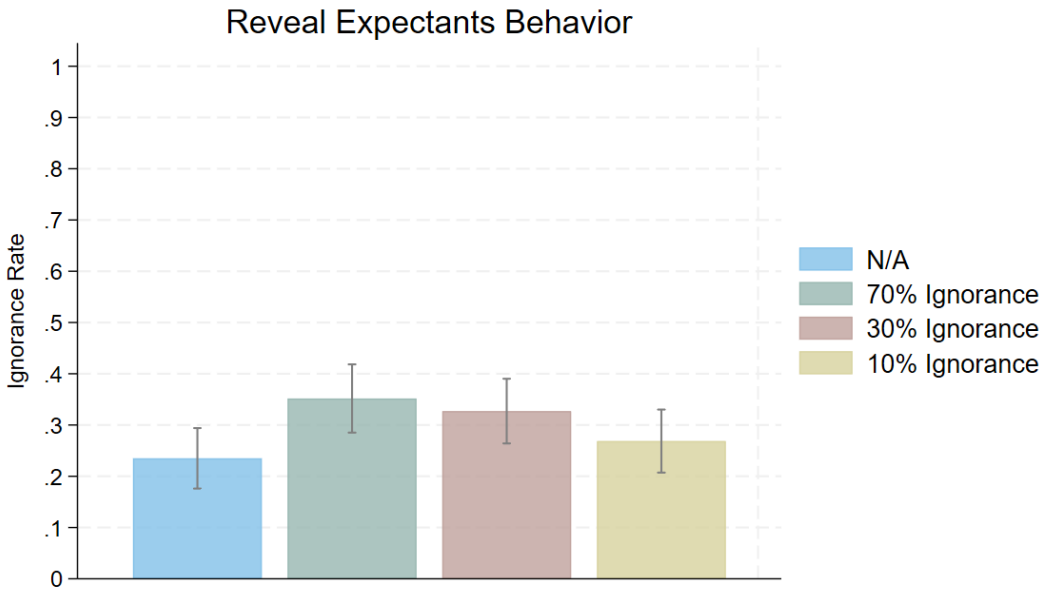


The results suggest that individuals adjust their behavior based on how much their prior beliefs deviate from observed norms. This effect is particularly pronounced among *ignorance expectant* dictators—those who initially believed that most others would choose to remain ignorant. Although they represent a minority of the sample (N = 206), these individuals responded strongly when exposed to a *10% ignorance* environment.[8] 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 of means test confirms this effect is statistically significant ($t = 2.73$, $p < 0.01$). 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.

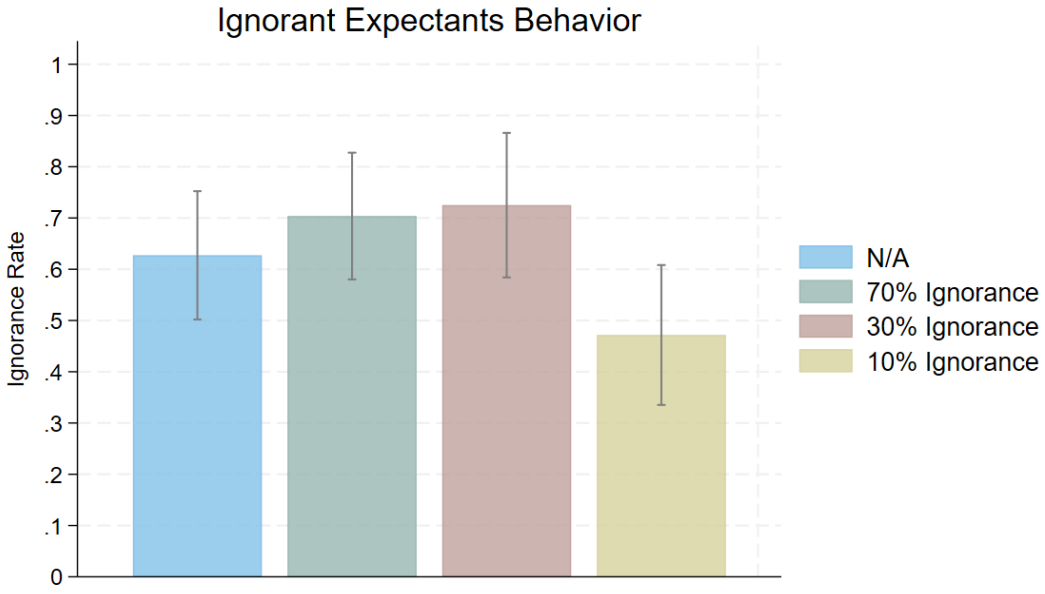
Result 2: Compared to *reveal expectant* dictators, *ignorance expectant*

[8]In Appendix A.5, 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: Reveal Expectants Versus Ignorance Expectants Dictators



(a) Ignorance rates of subjects who predicted an ignorance rate below 50%. N = 814.



(b) Ignorance rates of subjects who predicted an ignorance rate above 50%. N = 206.

Notes: Bars indicate 95% confidence intervals.

dictators were far less likely to reveal payoffs but significantly more likely to follow norms in the *10% Ignorance* environment.

In contrast, there is no support for *Hypothesis 3 - Exculpatory Beliefs*, 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. Instead, their information avoidance behavior remained largely unchanged regardless of the observed norm. These findings suggest that an ignorance norm can persist when groups of individuals reinforce each other's avoidance of information, creating localized clusters of strategic ignorance (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 ($D = 0.036, p = 0.897$). These 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 dictators made their allocation choice. Social appropriateness ratings, which are incentivized to match the modal response of other participants, reflect second-order beliefs about how others perceive the action. In contrast, personal appropriateness ratings capture individuals' own moral judgments, which remain consistent across conditions. Comparing across treatments, choosing not to reveal is rated as more socially appropriate in the *70% Ignorance* condition than in the *30% Ignorance* condition, suggesting that injunctive norms are updated in response to descriptive ones. However, personal appropriateness ratings are consistent across treatments.

This divergence is a departure from standard motivated cognition models, which would predict that individuals revise personal beliefs or moral assessments to align with self-interested behavior (Shepperd, Malone, and Sweeny, 2008). Instead, individuals appear to selectively update beliefs about injunctive norms, or how others view their actions, while keeping personal norms fixed. These findings, detailed further in Appendix A.6, provide no support for Hypothesis 4 (Self-Serving Beliefs).

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

Conclusion

This study demonstrates that individuals' willingness to avoid information is shaped by social norms, with lower information avoidance rates observed when norms favor transparency. Notably, *ignorance expectants*—those who assume that ignorance is widespread—adjust their behavior significantly when exposed to pro-information norms, whereas *reveal expectants* were more limited in their responsiveness to norm behaviors. Importantly, we find no evidence that individuals distort their beliefs to justify moral wiggle-room. Instead, behavior closely tracks the norm messages participants receive—suggesting that

individuals are genuinely aligning with perceived social expectations rather than engaging in motivated reasoning.

The experimental design implicitly treats information-seeking and information-avoidance as opposing normative social norms. That is, social environments can support either a transparency 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. These findings point to scalable policy interventions where small nudges can shift behavior at scale, especially in settings where stronger interventions are impractical. The findings also highlight that clusters of *ignorance expectant* individuals–those who believe that ignorance is the norm–are particularly responsive to updated information about prevailing norms, but only when those norms are overwhelmingly one-sided. This suggests that interventions targeting these groups with explicit messages about the widespread adoption of information-seeking behavior could be 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.

Interesting, ignorance rates were notably low even under the *No Norm* condition (32%) when compared to Mol, Soraperra, and Weele (2025) and Z. Grossman, Hua, et al. (2025) who reported ignorance rates of 60% and 62% respectively using a similar

Prolific subject pool. As all subjects had their beliefs elicited, simply asking subjects what they believe others will do may have significantly influenced behavior. Future research could investigate the direct effects of eliciting beliefs on norms.

Taken together, these findings suggest that strategic ignorance is not just an individual shortcut 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.

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.

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.

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 Perspectives* 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.

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.

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.

Weele, Joël J. van der, 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.

A Appendix

A.1 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 A.7.

Table 3: Subject demographics breakdown.

	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%)

A.2 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.

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.

A.3 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, Di-

mant, Gelfand, et al. (2023), in which individuals align their behavior with perceived descriptive norms when they believe those norms are both commonly followed and socially expected. Following **spiekermann2016objective**, 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-transparency 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-transparency* 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 transparency to be the norm (*transparency pooling*).

Such multiplicity is consistent with the experimental observation that *ignorance expectants* are particularly responsive to pro-transparency 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.

A.4 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.

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 4: 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 **??** shows the distribution of universalism scores.

As additional collaborative evidence, Table 7 shows that in the *Unknown* condition, political affiliation has no statistical effect, while in the *Known* condition, Republica-

tions, 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.

A.5 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 5.

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 to follow norms.

Table 5: Effect of Conformity Score on the Likelihood of Following Norms or Being a *Ignorance Expectant* Dictator

	(1) followed_norm	(2) followed_norm	(3) pessimistic	(4) pessimistic	(5) pessimistic	(6) pessimistic
conformity score	0.0548	0.0611	-0.0267	-0.0182	0.00576	0.00808
	(0.0421)	(0.0464)	(0.0572)	(0.0657)	(0.0550)	(0.0624)
Condition	All	All	Known	Known	Unknown	Unknown
Controls	No	Yes	No	Yes	No	Yes
Observations	675	654	464	429	455	423

Standard errors in parentheses

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

A.6 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.

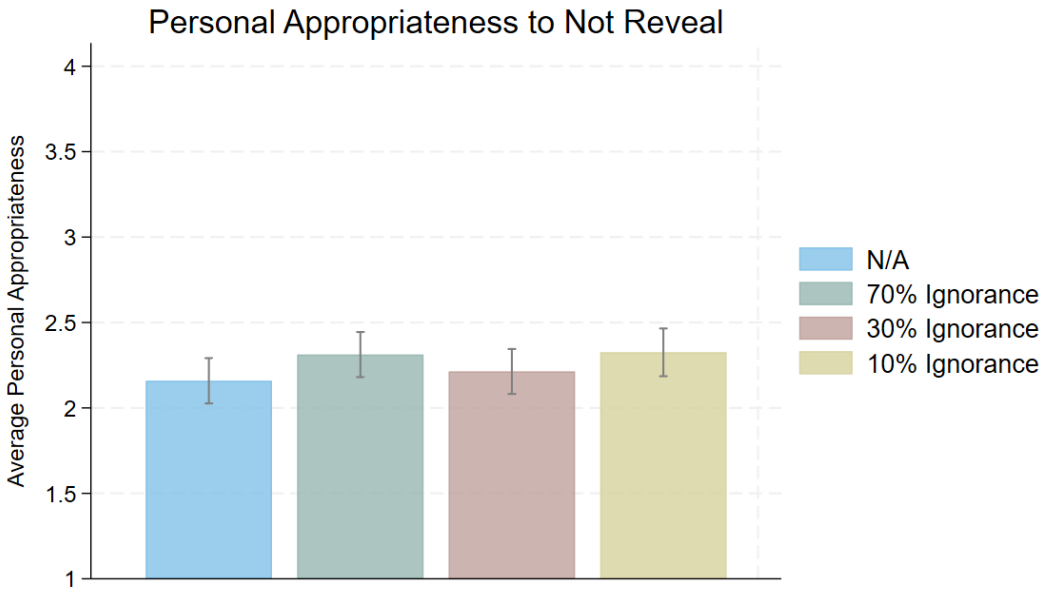
The following pre-registered hypothesis was tested:

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.

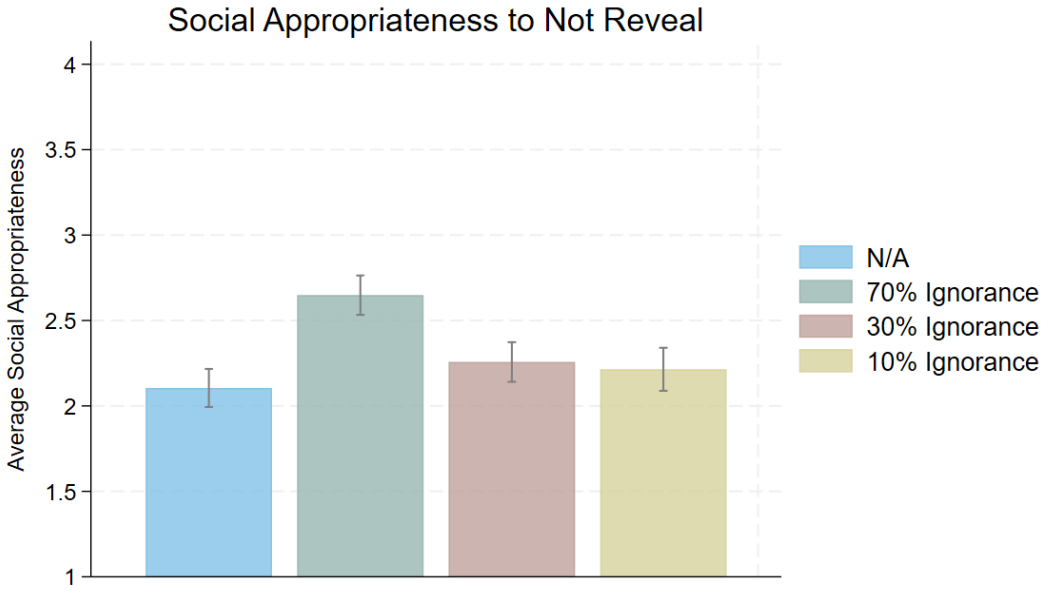
The results indicate that personal appropriateness ratings remained statistically consistent across all treatment conditions, suggesting that individuals' moral evaluations did not depend on their information status. In contrast, perceptions of others' social appropriateness judgments varied systematically, 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 5 presents these results, highlighting the distinction between personal appropriateness judgments and incentivized social appropriateness perceptions.

Figure 5: Social Appropriateness



(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.

Notes: Bars indicate 95% confidence intervals.

A.7 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 6 reports the regression results, which assess how subjects' demographic traits relate to their likelihood of avoiding information.

Table 6: 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

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

A.8 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 7 presents the full regression results.

Table 7: Demographics on Being a *Ignorance Expectant* Versus *Reveal Expectant* Dictator

	(1) pessimistic	(2) pessimistic	(3) pessimistic
pessimistic			
Age	-0.00167	-0.0126	0.0109
	(0.00513)	(0.00796)	(0.00732)
Female	0.00602	0.0595	-0.0364
	(0.105)	(0.158)	(0.155)
Black	0.203	0.408*	0.101
	(0.166)	(0.247)	(0.239)
Hispanic	-0.141	0.0798	-0.350
	(0.229)	(0.349)	(0.332)
East Asian	0.297	0.576*	-0.0216
	(0.241)	(0.349)	(0.375)
Southeast Asian	0.273	0.553	0.0408
	(0.291)	(0.390)	(0.489)
Other Race	-0.0859	0.245	-0.417
	(0.322)	(0.417)	(0.600)
less than $20,000	0.280	0.214	0.268
	(0.316)	(0.507)	(0.432)
less than $40,000	0.217	0.456	-0.187
	(0.335)	(0.528)	(0.465)
less than $60,000	0.0992	0.00115	0.146
	(0.337)	(0.538)	(0.457)
less than $80,000	-0.231	0.0480	-0.737
	(0.351)	(0.552)	(0.497)
less than $100,000	0.179	-0.0426	0.154
	(0.365)	(0.591)	(0.491)
greater than $100,000	0.150	0.126	0.0412
	(0.349)	(0.569)	(0.475)
Part-Time	-0.0436	-0.194	0.0578
	(0.158)	(0.233)	(0.242)
Self-Employed	-0.362*	-0.384	-0.355
	(0.188)	(0.267)	(0.291)
Unemployed	-0.0990	-0.402	0.0853
	(0.169)	(0.250)	(0.248)
Retired	-0.160	-0.158	-0.269
	(0.339)	(0.497)	(0.510)
Other Employment	-0.437*	-0.884**	-0.269
	(0.253)	(0.416)	(0.358)
High school graduate	-1.106	-0.361	-0.253
	(0.936)	(0.254)	(0.226)
Bachelors	-1.023	-0.338	-0.158
	(0.937)	(0.235)	(0.203)
Masters	-0.823	0	0
	(0.944)	(.)	(.)
Student	-0.0329	-0.193	0.211
	(0.174)	(0.271)	(0.245)
Independent	-0.0194	-0.342	0.243
	(0.143)	(0.222)	(0.206)
Democrat	-0.0292	-0.0419	0.0152
	(0.143)	(0.208)	(0.210)
Judaism	-0.318	-0.201	-0.479
	(0.425)	(0.670)	(0.598)
Hinduism	-0.187	-0.209	0
	(0.750)	(0.862)	(.)
Buddhism	-0.777	0	0.160
	(0.571)	(.)	(0.797)
Agnosticism	0.0333	0.0534	-0.0283
	(0.143)	(0.211)	(0.211)
Atheism	0.208	0.374*	0.0179
	(0.154)	(0.222)	(0.230)
Other Religion	0.206	0.582**	-0.142
	(0.167)	(0.250)	(0.248)
Condition	All	Known	Unknown
Observations	863	429	424

Standard errors in parentheses

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

A.9 Appropriateness Valuation of Revealing and Not Revealing Between *Ignorance Expectants* and *Reveal Expectants*

Table 8 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 both more personally and socially appropriate compared to *ignorance expectants*, who tend to assign higher appropriateness to not revealing. 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—independent of external coordination—play a central role in guiding economic behavior Bašić and Verrina, 2024.

	Ignorance Expectants	Reveal Expectants
Personal Appropriateness: Reveal	3.14	3.45
	(1.21)	(1.07)
Personal Appropriateness: Not Reveal	2.98	2.72
	(1.24)	(1.16)
Social Appropriateness: Reveal	3.32	3.58
	(1.15)	(1.03)
Social Appropriateness: Not Reveal	3.20	2.81
	(1.18)	(1.14)

Notes: Higher values indicate higher appropriateness ratings. Social appropriateness was incentivized by matching the modal response of other participants, while personal appropriateness was non-incentivized.

Table 8: Norm Appropriateness Ratings by Belief Type

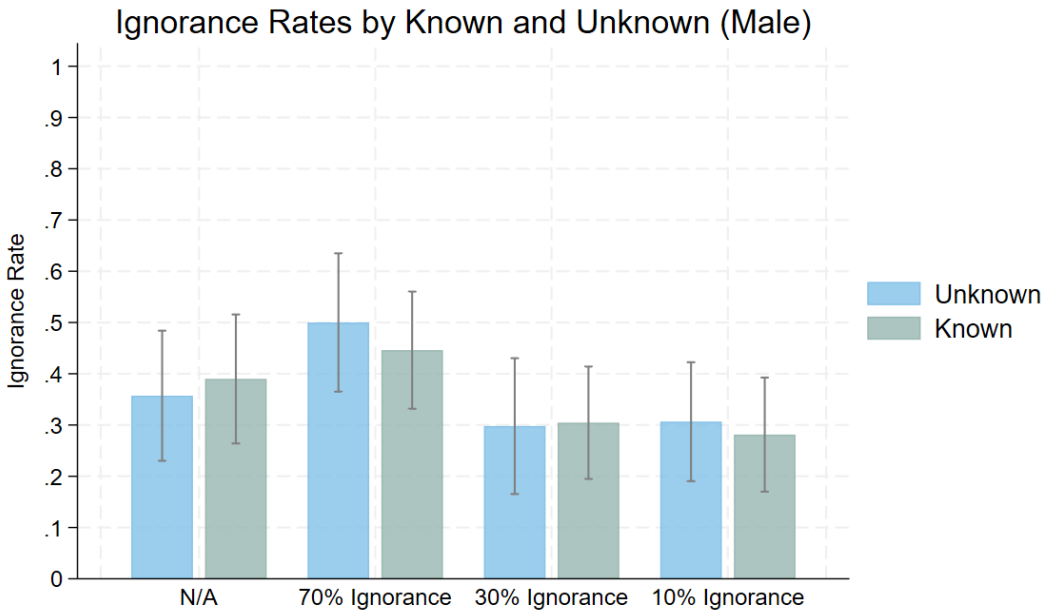
A.10 Exploratory Finding: Gender effects on Responses to Norms.

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 6 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.

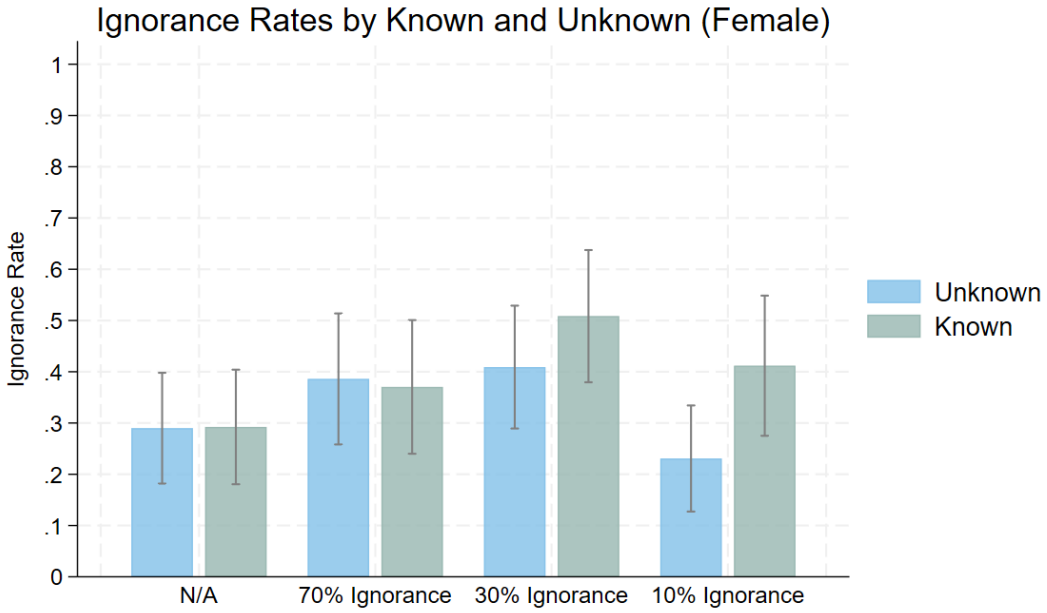
While previous work finds that women are more prosocial than men in dictator games (Eckel and P. J. Grossman, 1998), my results suggest that such gender differences are sensitive to the framing of social context. Specifically, when decisions are framed *self-referentially*, female participants exhibit 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 unconditional 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 in-

Figure 6: Framing and Information Interaction



(a) Ignorance rates of male subjects.



(b) Ignorance rates of female subjects.

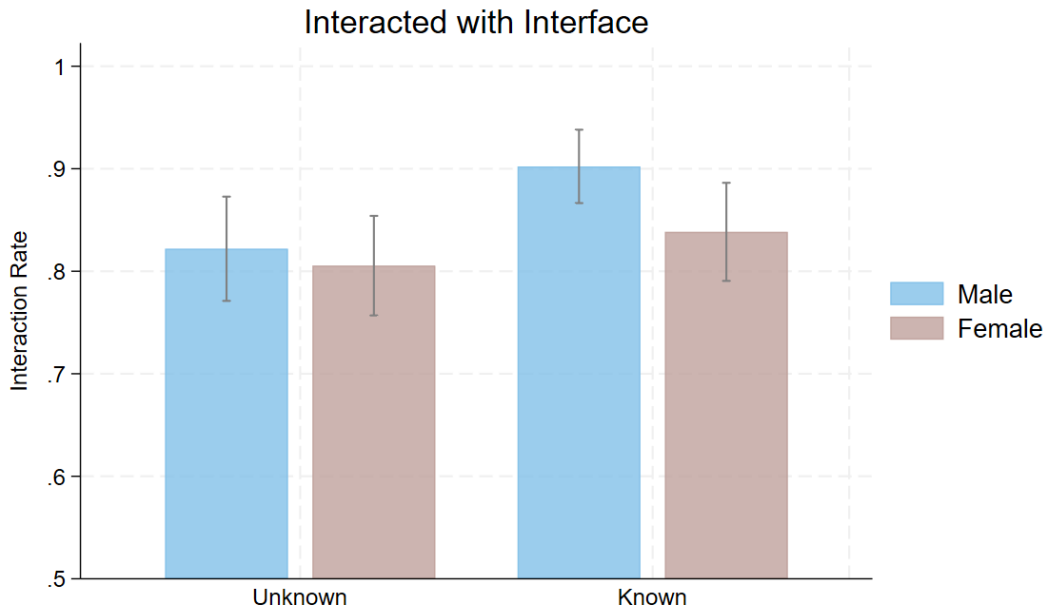
Notes: Bars indicate 95% confidence intervals.

teraction rate and time spent engaging with the mock-up interface between the *Known* and *Unknown* conditions. Figure 7 reports the choice process data. Under panel 7a, 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 7b, 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.[9]

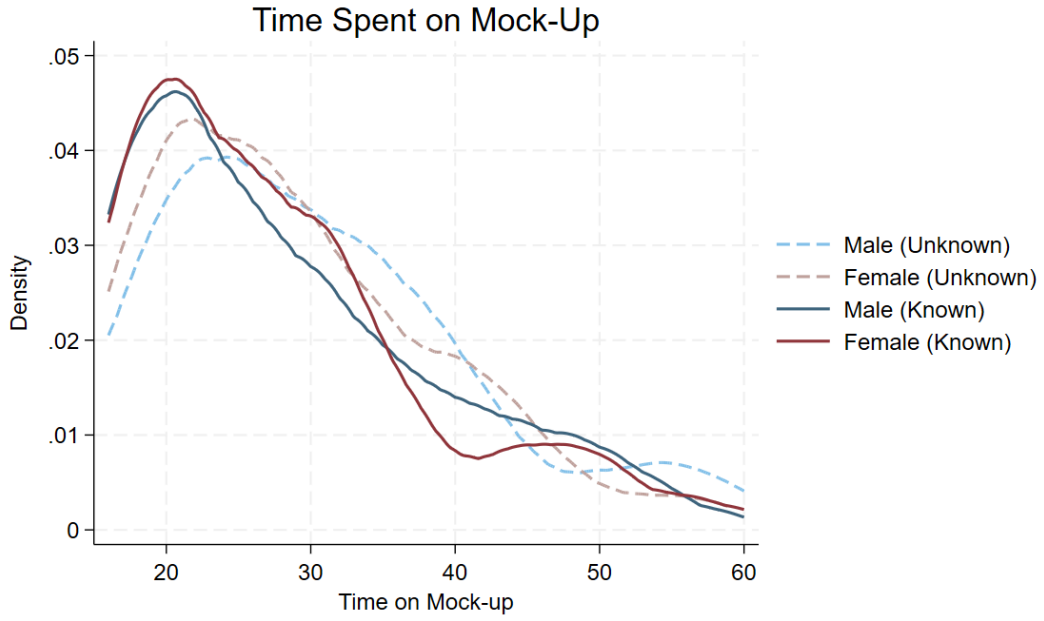
Examining subjects' moral universalism trust scores provides additional insight. In the *Known* condition, universalism negatively correlates with keeping the payoff tables hidden, whereas no such relationship emerges in the *Unknown* condition. This suggests that when individuals are explicitly aware 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 A.4.

[9] Appendix A.13 contains screenshots of the experimental interface.

Figure 7: Choice process metrics



(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.

Notes: Bars indicate 95% confidence intervals.

A.11 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.[10] In that study, dictators in treatment arms with nearly identical interfaces and instructions exhibited an average ignorance rate of 62%.[11] 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 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.[12]

The comparison highlights the significant impact of belief elicitation on reducing strategic ignorance, suggesting that prompting individuals to consider others' behavior can promote transparency, 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 willful

[10] 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.

[11] Appendix A.14 provides the interface details from Z. Grossman, Hua, et al. (2025).

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

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).

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 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.

A.12 Robustness Check for Result 5: Gender Differences in Norm Responsiveness

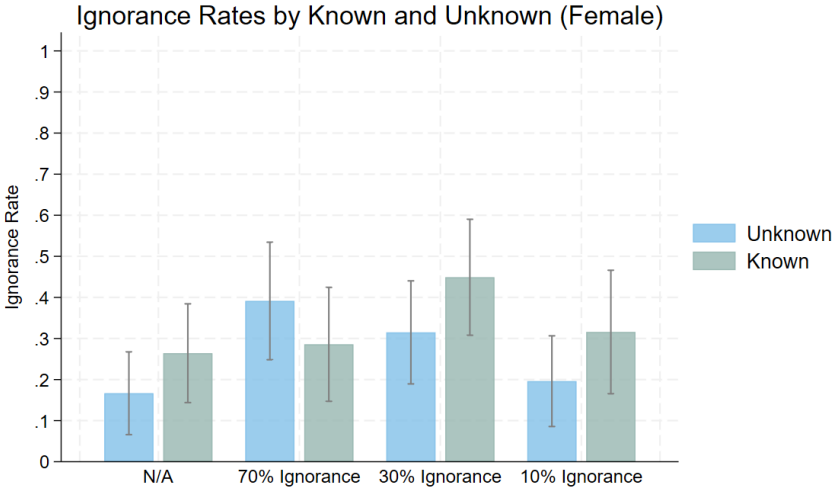
This appendix further evaluates Result 5 by focusing exclusively on *reveal expectant* dictators to isolate the effects observed in Result 2. By restricting the analysis to this subgroup, we examine whether the observed gender differences in norm responsiveness persist when *ignorance expectant* dictators—who exhibit stronger conformity to prevailing norms—are excluded.

The results indicate that women in the *Known* condition remain more resistant to normative pressures compared to men. This suggests that *self-centric* framing dampens the tendency for women to follow majority norms, whereas men exhibit greater norm responsiveness across conditions.

To formally test this effect, we conduct a two-sided difference of means test comparing the likelihood of following the majority norm across all three ignorance environments. The difference is statistically significant at the 5% level ($t = 2.24; p < 0.05$), reinforcing the conclusion that gender differences in norm adherence are influenced by framing effects rather than simply overall differences in beliefs.

Table 8 presents the full statistical results.

Figure 8: *Reveal Expectant* Female Ignorance Rates



A.13 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.

No one revealed | Everyone revealed

\- | Percentage of people who **revealed**: % | +

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

A minority: 30% or 3 out of 10 people chose to reveal the payoff tables.

You predicted that **51% of people revealed** the payoff tables.

You **overestimated** the amount of people who had revealed the payoff tables. People were more 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

No Norm treatment

You predicted that **24% of people revealed** the payoff tables.

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

Revealed payoffs

Actual payoff table shown:

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

Please select one of the options, then click to continue.

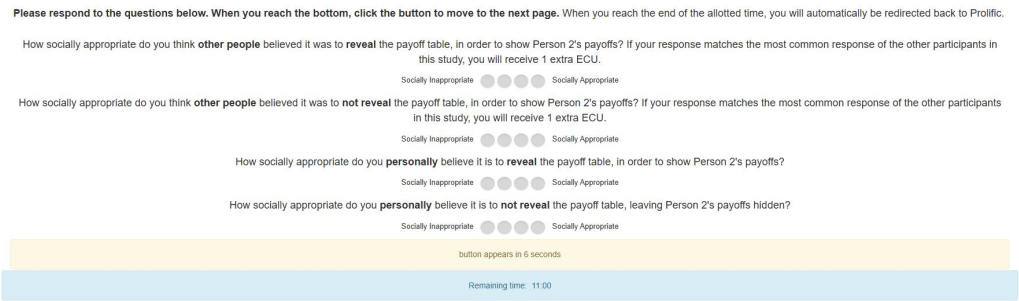
A

B

Continue

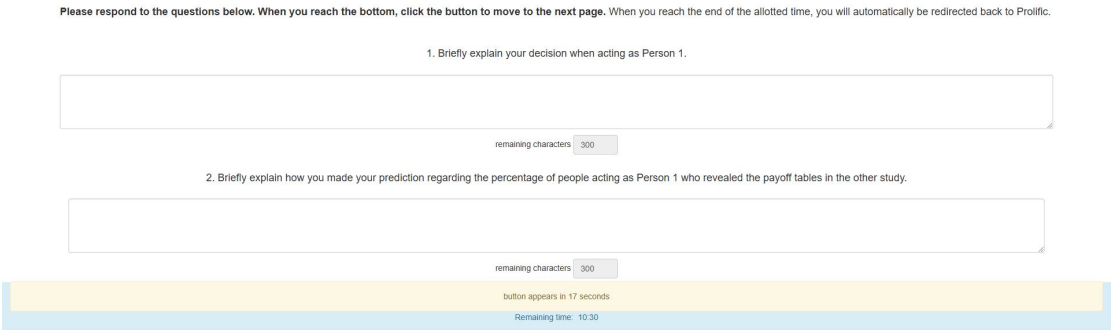
Page 8 Instructions

Appropriateness Rating



Page 9 Instructions

Free Response Explanation



Page 10 Instructions

The Conformity Scale

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.

I often rely on, and act upon, the advice of others.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I would like to be the last one to change my opinion in a heated argument on a controversial topic.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

Generally, I'd rather give in and go along for the sake of peace than struggle to have my way.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I tend to follow family tradition in making political decisions.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

Basically, my friends are the ones who decide what we do together.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

A charismatic and eloquent speaker can easily influence and change my ideas.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I am more independent than conforming in my ways.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

If someone is very persuasive, I tend to change my opinion and go along with them.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I don't give in to others easily.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I tend to rely on others when I have to make an important decision quickly.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

I prefer to make my own way in life rather than find a group I can follow.

Very strong disagreement ○ ○ ○ ○ ◉ ○ ○ ○ ○ Very strong agreement

Continue

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

11. How often do you go to church/temple or participate in any religious activities each month?

- 0
- 1
- 2
- 3
- 4
- 5
- 6+

Continue

Remaining time: 06:36

Page 12 Instructions

End Screen

Thank you for participating in this study! As an **additional bonus**, we will assign you as Person 2 (the recipient) for the decision made by a different Person 1. This bonus will be provided to you within 72 hours of the completion of this study.

Please keep this browser page open. Once the timer has elapsed, you will automatically be redirected to Prolific for completion and payment. **If you close the browser before the timer elapses, you will not be paid.**

Additional survey questions are on the following pages.

Continue

Remaining time: 05:56

Page 13 Instructions

Optional Moral Universalism Trust Scale

Please keep this browser page open. Once the timer has elapsed, you will automatically be redirected to Prolific for completion and payment. **If you close the browser before the timer elapses, you will not be paid.**

There are various ways in which people can cheat on others or take advantage of them. How much do you trust that different people will not cheat on you or take advantage of you? **In each row below, how would you split 100 "trust points" between a randomly-selected person who lives in the United States, and the individual displayed on the right (who is part of a particular social group)?** The closer you drag the slider to one individual, the more you trust that individual, relative to the other individual. **Please assume all of the individuals below live in the United States.**

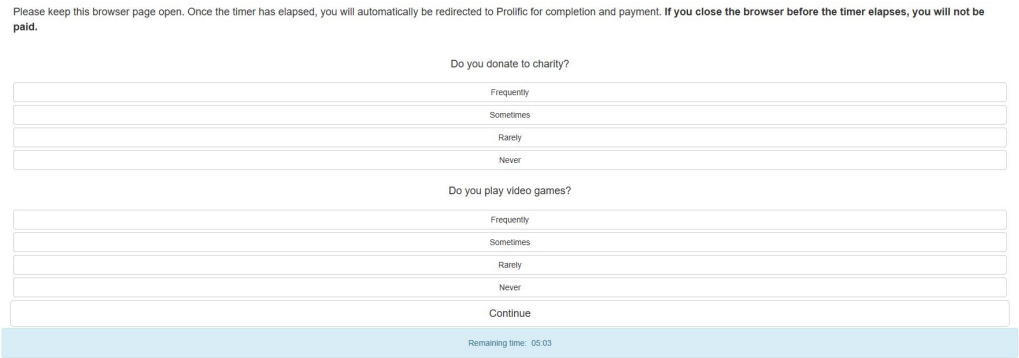
Randomly-Selected Person M who lives in the US	Trust this person -	Trust this person +	**Someone who shares your religious beliefs (e.g. a fellow Christian)**
Randomly-Selected Person N who lives in the US	Trust this person -	Trust this person +	**Someone of your same race/ethnicity (e.g. a fellow Hispanic person)**
Randomly-Selected Person O who lives in the US	Trust this person -	Trust this person +	**A member of one of your past current organizations (local church, leisure club or association, etc.)**
Randomly-Selected Person P who lives in the US	Trust this person -	Trust this person +	**Someone who lives in your local neighborhood**
Randomly-Selected Person Q who lives in the US	Trust this person -	Trust this person +	**A former or current colleague at work or school**

Continue

Remaining time: 05:35

Page 14 Instructions

Optional Questions



A.14 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