Whom do we learn from: Beliefs and Preferences

Akshay Moorthy*

November 28, 2025

Abstract

This paper examines how the identity of an information source shapes learning in settings where all stakeholders share the same incentives and information. The analysis draws on four large-scale online experiments designed to disentangle identity preferences from beliefs about the quality of an information source. The experiments are conducted with both naturally occurring identities (caste and religion in India) and experimentally assigned identities (in an EU/US sample). Across identity contexts, there is no evidence that preferences for the identity of an information source influence social learning. On the other hand, beliefs about information quality strongly influence learning, but participants are overconfident and often do not learn when it would benefit them. Finally, participants prefer to learn from a non-social source (a computer algorithm) rather than another human. The results highlight the importance of providing credible signals of information quality when social identities are salient, and offer new insights on how identity affects information acquisition.

Keywords: Social learning, social identity, discrimination
JEL Codes: D83, D91, J71

*I gratefully acknowledge guidance and support from Bertil Tungodden, Erik Ø. Sørensen, and Benjamin Enke. Kata Urban provided excellent administrative support. I further thank Kai Barron, Alexander W. Cappelen, Daniel Carvajal, Fabio Galeotti, Holger Herz, Sam Hirshman, Sissel Jensen, Michael Kosfeld, Julien Senn, and participants at various seminars and conferences for helpful comments. This study was supported by ERC grant 788433. The experiments were pre-registered at the AEA RCT registry (#0011066 and #0011924). All experiments were reviewed and approved by the Norwegian School of Economics Institutional Review Board (NHH-IRB 39/22 and 44/22). Moorthy: Faculty of Business and Economics, University of Lausanne, and FAIR – Norwegian School of Economics; akshayarun.moorthy@unil.ch.

1 Introduction

Social learning is fundamental to human evolutionary success (Boyd et al., 2011; Henrich, 2016) and influences many economic decisions (Mobius and Rosenblat, 2014). In social learning, decision-makers (DMs) extract information by observing others' actions and decide how to use that information. When DMs observe others' actions, they may also notice other attributes such as a person's race, religion, or gender – components of social identity – which may influence how they learn. The links between identity and learning have received attention, with research showing that people are more likely to respond to information from sources seen as closer to themselves or having similar political ideologies (Dolan et al., 2012; Maclean et al., 2019; Robbett et al., 2023; Garcia-Hombrados et al., 2024). However, little is known about the nature of this relationship – which is surprising given a large body of research studying how identity-based beliefs and preferences influence economic decision making in many domains (Tajfel and Turner, 1978; Akerlof and Kranton, 2000; Shayo, 2020; Charness and Chen, 2020). This paper addresses that gap by studying the mechanisms through which identity influences learning across a range of important identity contexts.

The conceptual framework for this study builds on decades of research on identity-based discrimination. When learning from others, DMs may use the source's identity to form beliefs about information quality, as in statistical discrimination (Phelps, 1972; Arrow, 1973). This suggests that identity salience could increase learning and improve decision-making when observing information quality is difficult. At the same time, DMs may have preferences for associating with or avoiding particular identity groups, as in taste-based discrimination (Becker, 1957; Guryan and Charles, 2013). This preference channel could lead to poorer decision-making, as individuals would ignore information not because of its quality but because of the source's identity. Such behaviour not only undermines effective communication but could also lead to belief polarisation or the formation of echo chambers at scale (Levy and Razin, 2019). Therefore, understanding these channels is crucial for designing communication strategies and information interventions targeted at improving decision-making. However, identifying preferences for the identity of information sources is empirically challenging as many determinants of information quality, such as income or education, are correlated with identity and may influence beliefs.

In this paper, I report on four incentivised online experiments (with $N = 2297$ participants) designed to disentangle beliefs about information quality from preferences for the information source's identity. The experiments share the same structure – a two-period task in which participants make incentivised decisions in both periods – and are deployed in different identity contexts. In the first period, participants make an independent decision in a cognitively demanding "balls-and-urns" task – a well-established paradigm for examining belief formation and learning (Benjamin, 2019; Weizsäcker, 2010) – where calculating the correct posterior probability is required to earn an incentive. In the second period, participants are shown a decision made by another individual – the *source* – on the same task. Participants then decide whether to learn from the source by *switching* to the source's decision, or not to learn by *sticking* to their independent decision.

This experimental framework focuses on an understudied type of learning in which decision-makers choose to learn from others despite having access to identical information. Both the decision-maker and the source make choices on the same task with the same information – information that is sufficient to yield an incentive-maximizing decision – and they operate under identical incentive structures. This setup provides a rational benchmark: fully informed Bayesian decision-makers would always make a correct decision in the first period and therefore have no reason to switch in the second. Consequently, a decision to switch in the second period cannot be attributed to a belief that the source possesses superior or different information – usually the focus of social learning experiments – but must instead reflect a belief that the source has a better interpretation of the same information. Understanding this form of learning is crucial because many real-world decisions involve seeking not only better information but also more effective interpretations such as when people solicit a "second opinion". For example, when choosing a health insurance or investment plan, DMs may choose to follow the decisions of their friends or peers despite having exhaustive information about the various options.

The online experiment setting provides control over the observable characteristics of the source and eliminates many channels that may influence learning indirectly (such as image or signalling concerns). This allows for disentangling the roles of beliefs and preferences, which is achieved by independently manipulating two attributes of the source. First, participants receive a randomly assigned signal of the information's *quality*. This is the probability that the shown source's decision is correct or, equivalently, the probability with which switching to the source would yield the incentive. Second, they observe the source's social *identity*, which is also randomly assigned and varies between treatment conditions in each experiment. As the source's quality is independent of the source's identity, comparing switching behaviour between various treatments reveals preferences for learning because of a difference in the source's identity.

Together, the four experiments provide a dataset of ≈ 13400 switching decisions made by 2297 participants. Two experiments study naturally-occurring identity contexts, caste and religion in India. These factors play a central role in Indian social and economic life (Munshi, 2019; Mosse, 2019; Iyer, 2016), and there are substantial differences in income and educational attainment between caste and religious groups (Census of India, 2011; Asher et al., 2024). A person's religion or caste is often salient as it can be accurately inferred from their name, mode of speech, body language, or other visible markers. In experiment *Caste*, the source belongs to either a high-status or low-status caste group. In experiment *Religion*, the source is either Hindu or Muslim. Caste and religious identities are made salient by assigning surnames to the source – a widely-used strategy deployed in many different contexts (Guryan and Charles, 2013; Bertrand and Duflo, 2017) including caste (Hoff et al., 2011; Sankaran et al., 2017) and religion (Fershtman and Gneezy, 2001; Chakravarty et al., 2016). As participants also have religious and caste identities, the design allows for the analysis of both general preferences – whether the DM learns more from a source of a particular caste or religion – and in-group preferences – whether the DM learns differently from a source sharing their religious or caste identity. Prior research indicates that identity also influences decision-making through in-group favouritism and out-group parochialism (Shayo, 2020), and the analysis examines these effects in detail.

Experiment *Minimal* focuses on in-group preferences using experimentally assigned identities. The experiment is based on the Minimal Identity paradigm (Tajfel and Turner, 1978; Chen and Li, 2009) – participants are randomly assigned to one of two minimal identity groups and see a source from either their in- or out-group. This experiment allows for the study of identity effects without invoking many of the empirical challenges associated with naturally occurring identities (Charness and Chen, 2020). Finally, experiment *Human vs. Computer* compares learning from social versus non-social sources by presenting participants with decisions made either by a human or by a computer algorithm. This experiment tests whether individuals learn differently from algorithmic recommendations than from human sources and also tests the effectiveness of the identity manipulation technique used in the experiments. Participants in the *Minimal* and *Human vs. Computer* experiments are recruited via an online labour platform whose participants are from the EU/US.

The first main finding is that a considerable fraction of participants switch to the source, showing that participants choose to learn even though sources have the same information. The analysis shows that participants are responsive to the randomly assigned quality signal: the propensity to switch increases with an increase in information quality. However, most participants leave money on the table by not switching even when the source is high-quality. This behaviour is partially driven by participants' overconfidence – more confident participants learn less, and overall, individuals overestimate the accuracy of their initial decisions. These patterns are qualitatively consistent across all samples and identity contexts, though responsiveness to information quality varies. The analyses show parallels to learning in decision contexts where information sets differ, and provide suggestive evidence of asymmetric learning.

The second main finding is that people do not express a preference for learning from a particular identity in the *Caste*, *Religion*, and *Minimal* experiments. Participants are equally likely to switch to a source from a high- or low-status caste group in experiment *Caste*, a Hindu or a Muslim source in experiment *Religion*, or an in-group or out-group source in experiment *Minimal*. At the same time, there are substantial differences in underlying beliefs about the performance of different identity groups on the balls-and-urns task: high-status caste individuals are perceived to outperform low-status caste individuals, and Hindus are perceived to outperform Muslims.

Additional analyses using self-reported religious and caste identities reveal no differences in switching behaviour when decision-makers share the source's identity compared to when they do not. Several robustness tests and heterogeneity analyses – examining biases in belief updating, exposure to religious and caste diversity, and attitudes toward caste-based affirmative action – further support these null results. Although there is weak evidence of preferences for certain overlaps between religious and caste identities, overall, the results strongly suggest that participants do not have a systematic preference for the identity of information sources when those sources are human, their identities are salient, and their quality is known.

The third main result is that people prefer to learn from a non-social source (a computer algorithm) rather than from a social source (an anonymous human). Participants in experiment *Human vs. Computer* are 15% points more likely to switch to the source when the source is a computer than when it is a human. The effects are quantitatively large (about 0.3 standard deviations) and are the same at both low and high levels of information quality. This result supports the idea of "algorithmic appreciation" – a general preference for learning from a non-social source than from another human – that has been the subject of investigation in the field of human-computer interaction (Logg et al., 2019; Hou and Jung, 2021).

This paper contributes to the empirical social learning literature (reviewed recently by Mobius and Rosenblat (2014)), in which the role of the identity of information sources has been largely overlooked. First, the results show that beliefs about information quality strongly influence learning, while preferences for specific identities or for their in-groups do not appear to play a significant role. This finding complements experimental research documenting identity-based differences in information processing in a variety of settings: gender differences in intra-household social learning (Conlon et al., 2021), political in-group biases (Robbett et al., 2023; Zhang and Rand, 2023), national in-group following (Dekel and Shayo, 2023), within the minimal identity paradigm (Berger et al., 2018; Parys and Ash, 2018; Zou and Xu, 2022), and agricultural technology adoption (BenYishay and Mobarak, 2018). The results also relate to research on the messenger effect, which finds that information transmission and belief formation are affected by the information source's identity (Dolan et al., 2012; Maclean et al., 2019; Wabitsch, 2024; Garcia-Hombrados et al., 2024; Afrouzi et al., 2024). Relative to these works, this paper focuses on identifying whether identity-biased information acquisition or belief updating is driven by beliefs or preferences related to identity. The prominence of the belief channel (relative to the preference channel) is consistent with findings on information quality (Robbett et al., 2023) and perceived knowledge (Dekel and Shayo, 2023). The results contrast, however, with those of Bauer et al. (2023), who find evidence of preferences for information from political in-groups.

Second, the experiments in this paper provide a novel perspective on learning by focusing on situations where all stakeholders have the same information and incentives, and shed light on understudied situations such as the demand for second opinions when making high-stakes decisions. The results show that a sizeable share of people choose to learn from others even though a rational benchmark suggests that switching is never beneficial. While this type of situation shares some features with studies on conformity, the key difference is that there is no role for social interactions as decisions are unobservable by other people. The analyses also indicate that several mechanisms that affect learning when information sets differ (Weizsäcker, 2010; Mobius and Rosenblat, 2014; Benjamin, 2019; Barron, 2021) also exist in this setting. These patterns highlight a potentially powerful source of learning where people seek better interpretations of the same information, a mechanism that can influence learning even when information sets differ.

This paper also speaks to research on the economics of identity, in particular to the literature focusing on social identity in India, by providing direct evidence of the influence of religion and caste on learning. Caste and religion are important aspects of social life and economic behaviour in India (see Iyer (2016); Munshi (2019); Mosse (2019) for reviews), and a large body of prior research has documented the influence of religious and caste identity in various decision-making domains such as consumption (Atkin et al., 2021), hiring in labour markets (Siddique, 2011), labour supply (Cassan et al., 2019; Oh, 2023), marriage markets (Banerjee et al., 2013), and teamwork (Ghosh, 2022). This paper studies the role of caste and religion on learning, and shows that preferences for the religious or caste identity of information sources may not affect learning when the quality of information is salient. The results also show systematic differences in beliefs about the abilities of different caste and religious groups, and suggest that these beliefs may be used to infer information quality when the latter is not observable.

Finally, this paper contributes to research on human-computer interactions by comparing social and non-social learning. This literature has previously found mixed results: while some have found that people are averse to learning from algorithms and favour learning from humans (Goeree and Yariv, 2015; Dietvorst et al., 2018, 2015), others have found that people display algorithm appreciation by favouring algorithms over humans in certain situations (Logg et al., 2019). More recent research shows that these differences may depend on contextual features of decision tasks (Hou and Jung, 2021) or nomenclature (Langer et al., 2022). By holding information sets constant between both the human and computer sources, the design allows for a clean comparison of whether the nature of the source matters. The results show support for algorithm appreciation, suggesting that when information quality is precisely known, people prefer to learn from a non-social computer algorithm than from another human.

The paper is organised as follows: Section 2 presents the experiment design, provides intuition for the identification and hypotheses, and describes the various experiments and identity contexts. Section 3 examines the causal role of beliefs using the experimental variation in source quality. Section 4 focuses on causal evidence of preferences on learning, supported by robustness tests and heterogeneity analyses. Section 5 concludes with a discussion of the results and open questions for future research.

2 Experiment Design

The study consists of four main experiments, each with two treatment conditions. Participants in an experiment are randomly assigned upon entry to one of two treatment conditions which differ only in the identity of the source. All experiments use the same task structure, but differ in the identities used and the manner in which the identities are made salient.

This section starts with a description of the common experiment structure, followed by a conceptual discussion, and then presents the details of the different identity contexts of each experiment along with relevant background information. A complete set of screenshots of the experiment are presented in Appendix D.

2.1 Experimental task

The experimental task builds on the "balls-and-urns" paradigm, which is widely used to study belief formation and social learning. Participants complete six tasks where each task takes place in two periods and participants make an incentivised decision in each period.

First period. Participants are shown two urns, urn A and urn B, which contain 100 red or black coloured balls. Urn A contains θ red balls, and urn B contains $100-\theta$ red balls. Urn A is randomly selected with probability p and the participant does not know which urn is selected. k balls are drawn (with replacement) from the chosen urn and shown to the participant. Participants use a slider to make an incentivised decision $y_1 \in (0, 100)$ of the probability that urn A was chosen. After making the decision, participants state their confidence that y_1 is within ±2% of the correct decision. The tasks are pre-defined in the sense that the values of the base rates p, urn composition θ, and the signals (colours of k drawn balls) are the same for all participants. This design choice ensures that participants are exposed to tasks of the same difficulty and eliminates several practical challenges. Participants see tasks in a randomised order. The tasks, corresponding parameters, and other details are listed in Appendix Table B.1.

Participants can calculate the correct answer for each task by applying Bayes' rule using the provided information.[1] Participants are aware of this as they are given detailed instructions and work through a training task on how to use Bayes' rule to calculate the correct answer. In free-text feedback, many participants indicated that they tried to apply Bayes' rule while working on the task. However, the task is cognitively demanding which means that learning from others may be beneficial. Participants have an opportunity to do so immediately after they make their first-period decision.

Second period and main outcome. Participants are shown a decision $y_s \in (0, 100)$ made by another person, the source s, on the same task. Participants know that the source had the same information and saw the same draws. Participants see (i) the group *identity* g of the source, which is randomly assigned between-subjects, and (ii) a signal Q of the *quality* of the decision. Figure 1 shows the representation of these elements in the experiment interface. Next, participants make an incentivised binary choice y_2: whether to *stick* to their first decision ($y_2 = y_1$) or *switch* to the shown decision ($y_2 = y_s$). This is the main outcome variable, *Switch*, which is 1 when $y_2 = y_s$ and 0 when $y_2 = y_1$.

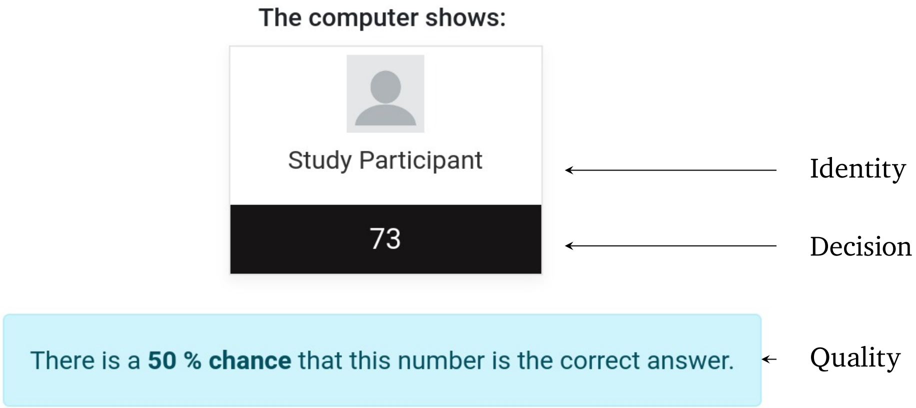
Eliciting learning behaviour using this discrete outcome rather than a conventional continuous outcome (over the belief space) has two advantages. First, the binary outcome creates a stark distinction between updates that are driven by a choice to move away from one's initial decision, and updates that reflect some degree of hedging (due to the uncertainty induced by the complexity of the balls-and-urns problem).[2] Second, the framing helps make switching psychologically meaningful and is much simpler for participants to understand – a key consideration when deploying this experiment in multiple contexts.

Source Decisions. Behind the scenes, each task has two source decisions associated with it. One of these decisions is correct and the other is incorrect. These decisions were extracted from the incentivised first-period decisions made by participants in separate experiments (conducted with an Indian sample and on Prolific). The decisions were chosen to satisfy a few criteria, the most important of which was that each of these decisions was made by at least one person from each identity group. Table B.1 provides a list of the correct and incorrect source decisions for each task. Additional details on how the sources are curated are provided in the discussion of the separate experiments in Section 2.3 below and in Appendix C.3.

[1]The correct answer can be calculated as follows: Let the number of drawn red balls $= r$. Then, $P(\text{draws}|A) = \binom{100}{k} \cdot \frac{\theta}{100}^{r} \cdot \frac{100-\theta}{100}^{k-r}$, $P(\text{draws}|B) = \binom{100}{k} \cdot \frac{100-\theta}{100}^{r} \cdot \frac{\theta}{100}^{k-r}$, and finally, $y_T = P(A|\text{draws}) = \frac{p \cdot P(\text{draws}|A)}{p \cdot P(\text{draws}|A) + (1-p) \cdot P(\text{draws}|B)}$. At the time of data collection, LLM tools such as ChatGPT were not good at solving this task.

[2]While this design choice prevents participants from altering their estimates partially, it shifts the focus onto beliefs and preferences regarding the source which is the focus of the experiment.

Figure 1: Source – Attributes



Notes. The figure shows how the source is represented in the experiments. *Identity* is randomly assigned between-subjects within each experiment group. The *quality* of the source's decision is randomly assigned at the task-level, and is the probability with which the shown *decision* is objectively correct ($= y_T$).

Quality signal. Participants are given a signal of the quality Q of the source's decision which is the probability that y_s is the correct decision. This means that with probability Q participants are shown a correct decision, and with probability $100 - Q$, participants are shown an incorrect decision. The quality varies within-subject and is chosen randomly for each task: $Q \in \{0.5, 0.9\}$ in *Caste*, or $Q \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ in the other experiments. This signal provides participants with precise beliefs about the quality of the decision that they could choose to learn from, and achieves the objective of making information quality independent of the source's identity. Additional details on these signals and a discussion of the rationale behind structuring the signals in this manner are provided in Appendix C.4.

Incentives. Participants complete six tasks, and make two decisions in each task. Each task has different configurations of balls, urns, and associated probabilities. One of the 12 decisions is randomly chosen for a bonus reward. Participants get a $3 bonus payment if the chosen decision is within ±2% of the true posterior; otherwise, they get no bonus. This incentive scheme was chosen both for simplicity, and to avoid skewing participant decisions as highlighted in recent literature (Danz et al., 2022). All participants received a flat $2 fee for participating in the study. Participants were paid in cash or in "channel points" with a value equivalent to the stated dollar amounts.

2.2 Conceptual framework and identification

In these tasks, a decision maker (DM) must choose the correct decision y_T in order to earn an incentive. Although the DM has sufficient information to identify y_T, the cognitive demands of the task mean that a non-Bayesian DM may find switching to the source's decision beneficial. In contrast, a fully Bayesian DM would identify the correct decision immediately and would have no incentive to switch.

The DM's decision to switch is a function of their beliefs about the quality of the information provided by the source and their confidence in the accuracy of their initial decision. The DM would be more likely to switch if they believe that the source provides high quality information, and less likely to switch the more confident they are about their independent decision's accuracy. At the same time, switching might involve psychological costs which would cause DMs to prefer their initial decisions. Prior research supports this view, showing that individuals tend to update their beliefs conservatively (Benjamin, 2019) and resist incorporating new information (Weizsäcker, 2010; Conlon et al., 2022).

Therefore, DMs switch if the utility from switching exceeds the utility from sticking:

$$U(E[y_2 = y_s], k) > U(E[y_2 = y_1])$$

where k represents the cost of switching or a preference for sticking to the initial decision. In the experiment, the expected values of switching or sticking are only affected by DM's beliefs about the quality of the source's information ($Q = E[y_2 = y_s]$) and their own decision's accuracy ($E[y_2 = y_1]$). This framework generates two empirically verifiable predictions relating to beliefs:

Prediction.

1. *Switching increases with an increase in beliefs about information quality (Q).*
2. *The more confident DMs are about the accuracy of their first decision (y_1), the less likely they are to switch to the source.*

DMs also observe the identity, g, of the source, which can affect these beliefs and preferences. First, the DM may make inferences about the source's task-relevant abilities based on perceived characteristics of the source's social groups, in a form of statistical discrimination (Phelps, 1972; Arrow, 1973). If a DM holds prior beliefs about the ability associated with a particular identity group, these beliefs will also influence their assessment of the source's decision quality. Consequently, DMs respond differently to sources of varying identities, driven by differences in these group-related beliefs. If these beliefs are incorrect, the DM's behaviour could resemble inaccurate statistical discrimination (Bohren et al., 2023).

Second, the DM may have a preference for (or aversion against) learning from sources of specific identities. Building on theories of taste-based discrimination and social identity research (Becker, 1957; Guryan and Charles, 2013; Akerlof and Kranton, 2000), which posit that membership to an identity group motivates decision-making to align with identity-driven behavioural prescriptions, the salience of social identity might lead individuals to ignore information not because of its quality but because of the identity of the source. In other words, the cost (preference) of switching k_g may vary based on the source's identity. If the DM experiences aversion or identity-driven social pressures against associating with a particular group, switching becomes relatively costly. Conversely, affinity for a particular group reduces the cost of switching.

The goal of the experiment is to causally identify whether k varies based on the source's identity. The main empirical challenge is that both preferences for switching to sources of particular groups and beliefs about groups jointly determine whether the DM chooses to switch. Thus, to identify preferences for the identity of information source, beliefs about the quality of information must be

Table 1: Experiment overview

Experiment	Identity Treatment	Sample	Sample size	Identity salience	Dates
Religion	*H* – Hindu	Panel survey	431	Surnames	April 2023
	M – Muslim		422		
Caste	*G* – General	Panel survey	415	Surnames	September 2023
	O – SC/ST/OBC		436		
Minimal identity	In-group	Prolific	142	Klee-Kandinsky	March 2023
	Out-group		136		
Human vs. Computer	Computer	Prolific	163	Labelling	March 2023
	Human		430		

Notes. Participants are assigned to one experiment, and to one identity condition within each experiment. General – Source is from the General caste category. SC/ST/OBC – Source is from the Scheduled Castes, Scheduled Tribes, or Other Backward Classes categories. In treatment *Human*, the sample from the *Minimal* experiments is also used. Appendix Table B.2 provides information on the demographics of the different samples.

controlled so that they are independent of identity. This is achieved through the random assignment of the quality signal Q, which breaks the link between beliefs about information quality and identity. Because this signal is precise and exogenously assigned, any observed differences in switching because of a difference in the source's identity can be attributed to preferences (or different costs) for learning from one group relative to another.

2.3 Identity contexts

Participants in an experiment are randomly assigned upon entry to one of two treatment conditions which differ only in the identity of the source. Table 1 provides a list of experiments in this study. Appendix Table B.2 provides summary statistics of the demographics of the different samples.

Experiment *Caste*

Background information. Caste is a system of social stratification in India that consists of thousands of caste groups called *jatis*. A person's caste is largely determined by birth. Many prominent features of the caste system such as endogamy, social hierarchy, segregation, and ritual purity continue to influence modern Indian society (Mosse, 2019; Munshi, 2019). Caste groups are often associated with occupations (Cassan et al., 2019) – for example, some castes are perceived as intellectuals, some as entrepreneurial or business-oriented, and others as farmers, etc. Some specific occupations are associated with population groups which were historically excluded from the caste system and regarded as "untouchables". While the caste system is largely linked with Hinduism, non-Hindus may also hold caste identities. To effectively target welfare programs and affirmative action policies, the Government of India classifies castes into four "categories". The "General" category (also known as the Forward Castes, ≈ 30% of the population) consists of *jatis* that are considered socially and economically advanced (in relative terms). The "Scheduled Castes" (SC, ≈ 20%) and "Scheduled Tribes" (ST, ≈ 9%) categories are formed of castes that are the most economically and socially disadvantaged. This category also contains the erstwhile "untouchables" or *Dalits*, and people from indigenous tribes. Last, the "Other Backward Classes" (OBC, ≈ 40%) category contains many *jatis* that are economically and socially disadvantaged relative to the General category.

Caste identity can influence both beliefs and preferences about information quality. First, individuals may hold beliefs about the cognitive ability of people belonging to particular caste groups. This could be driven by economic realities such as disparities in education and income between caste groups, or by caste-specific stereotypes related to occupational and educational choices. Second, people have strong preferences for associating with others from their caste groups or for avoiding associating with caste groups considered lower in status. Such caste-dependent preferences could impact how people learn from sources belonging to different caste groups.

Implementation. The caste identity of the source is made salient through surnames that are informative of the caste category of the individual. These surnames are shown along with an arbitrarily chosen initial (for example, Mr A. Moorthy) in place of "Study Participant" in Figure 1. There are two treatments in this experiment: G and O. In treatment G, surnames belonging to the General (or "Forward Castes") category are used. In treatment O, surnames belonging to the Scheduled Castes, Scheduled Tribes, or Other Backward Classes (SC/ST/OBC) are used. The surnames were validated for recognisability through a separate survey conducted through the same provider.[3] A list of surnames and additional details are presented in Appendix C.

The *Caste* experiment and the validation study were conducted through online surveys with a gender-balanced sample of Indian Hindu participants recruited through an online panel survey provider. The effective sample size for the main experiment is 851 Hindu participants. ≈ 63% of the participants belonged to the General (G) caste, and the remaining ≈ 37% belonged to one of the other caste groups. The G category is over-represented in the sample relative to the general population.

A supplementary experiment, *Caste – No Signal*, was conducted at the same time as the *Caste* experiment. The only difference between these two experiments was that participants in *Caste – No Signal* did not receive a quality signal in the second period.

Experiment *Religion*

Background information. Religion is an important part of daily life in India.[4] ≈ 80% of the country is Hindu, ≈ 14% is Muslim, and the rest of the population follows other religions such as Christianity, Sikhism, Buddhism, Jainism, etc (Census of India, 2011). Religion moderates social and market interactions, and inter-religious relations are often characterised by a lack of trust. In particular, tensions between Hindus (the majority religion) and Muslims (the largest minority) have deep historical roots which were exacerbated during the British colonial period and the eventual Partition of India in 1947. Contemporary conflicts arise from a variety of factors such as territorial disputes, religious nationalism, political polarisation, and competition for resources (Iyer, 2016; Iyer

[3] ≈ 350 individuals were incentivised to correctly classify a list of surnames into one of the four caste categories.

[4] Two-thirds our sample say that religion is a very important part of their daily lives, and 90% say that it is either somewhat or very important. This echoes findings from other surveys like the Gallup World Poll and the World Values Survey.

and Shrivastava, 2018; Jaffrelot, 2021). A recent study by the Pew Research Center (2021) finds that ≈ 21% of Muslims and 17% of Hindus report having experienced discrimination because of their religion in the past 12 months.

Muslims are, on average, poorer and less educated than Hindus (Asher et al., 2024), which may lead to beliefs that Muslims are less likely to be well-educated than upper-caste Hindus and therefore less likely to perform well on the balls-and-urns tasks. In terms of preferences, research has provided evidence of taste-based discrimination in the labour market (Thorat and Attewell, 2007), and that religious identity affects trust and pro-sociality (Dhami et al., 2024). Thus, it is plausible that religion can affect social learning through both belief and preference channels.

Implementation. The religious identity of the source is made salient using surnames that are informative of the religion of the individual. In treatment *Hindu*, the names are the same as those used in the *G* treatment of the *Caste* experiment. In treatment *Muslim*, the surnames are common Muslim surnames. The Muslim surnames were not validated separately as these are generally easily identifiable by Indians. A list of surnames and additional details are presented in Appendix C.

This experiment was conducted through online surveys with a sample of participants from India provided by an online panel survey provider. The effective sample size is 853 participants, of whom 647 were Hindus (≈ 75%), 67 Muslims (≈ 8%), and the rest (≈ 17%) of other religions. Muslims are under-represented in the sample relative to the population.

Experiment *Minimal Identity*

Background information. The minimal identity paradigm has been used for decades by social scientists to study the effects of identity on various decisions (Tajfel et al., 1971; Chen and Li, 2009). Participants begin by making choices in an innocuous task, often choosing between paintings by the artists Wassily Kandinsky and Paul Klee. These choices are used to classify participants into arbitrary groups unrelated to naturally occurring identities. Research has demonstrated that this method induces people to behave in a "groupy" manner by making choices that favour their in-group over their out-group (see Charness and Chen (2020)). The method has been used to study in-group effects in learning, although the focus has been on learning in settings with private information: for example, Berger et al. (2018), Parys and Ash (2018), and Zou and Xu (2022) find that people learn more from their in-groups in experimentally assigned identity settings. An unanswered question is whether these documented effects are driven by beliefs or preferences.

Implementation. In experiment *Minimal*, I follow the method used in Chen and Li (2009) to assign minimal identities to participants. First, participants are asked to examine a few pairs of paintings and indicate which they preferred in each pair.[5] Next, participants are classified into either the Orange or Purple group based on whether they liked Klee (Orange) or Kandinsky (Purple) paintings. Participants are informed of the method by which the groups were assigned. Participants are also

[5]One of the paintings in each pair was created by the artist Paul Klee, and the other by the artist Wassily Kandinsky. Paintings by these artists have been used in many studies that use this method because of their similarity (at least, to the untrained eye).

reminded of their assigned group identity halfway through the six tasks. Participants see sources as in Figure 1 with the addition of an orange or purple coloured label with the text "Orange group" or "Purple group", which depends on the participant's group and the treatment to which they are assigned. In the *In-group* treatment, they see sources belonging to the same group and in *Out-group* they see sources from the other group.

Participants in this experiment were recruited through Prolific (a popular survey platform for social science experiments) and came from the US, UK, and EU. The effective sample size is 278.

Experiment *Human vs. Computer*

Background information. Understanding human-computer interaction has become increasingly important with developments in generative artificial intelligence and the general importance of the internet as a medium in all aspects of social and economic life. However, the evidence on whether people learn more from humans or computers is mixed. While some find that people are more averse to learning from algorithms than from other humans (Goeree and Yariv, 2015; Dietvorst et al., 2018, 2015), others find evidence of algorithm appreciation (Logg et al., 2019). Hou and Jung (2021) show that appreciation or aversion emerges because of differences in framing and perceived accuracy, and Langer et al. (2022) document the sensitivity of people's responses to the terminology used to refer to the non-social source. The *Human v. Computer* experiment studies whether differences in how people learn from humans and algorithmic sources is driven by preferences for non-social sources. This experiment also had the objective of validating the identity manipulation technique (cosmetic and text modifications to the user interface) used in the other experiments.

Implementation. In this experiment, participants see either an anonymous human or a "Computer". In treatment *Computer*, the source is labelled as "Computer", and the human icon is replaced with a computer icon. Similar to the other treatments where estimates made by other humans are shown, participants are told that the signal indicates the probability that the shown estimate is correct, and that a randomly chosen number is displayed otherwise.[6] The comparison group for this treatment are participants pooled from a separate treatment within this experiment (*Human*, where sources are labelled as "Study Participant" as in Figure 1) and responses from the minimal group experiment.[7]

Participants in this experiment were recruited through Prolific, and the effective sample size is 593.

Procedural details

The experiments were programmed using OTree (Chen et al., 2016). The experiments on Prolific were conducted in March 2023. The *Religion* experiment was conducted in collaboration with Faktum Research in April 2023. The *Caste* experiments were conducted in September 2023 in collaboration with Norstat. The experiments were pre-registered at the AEA RCT registry (#0011066 and

[6] This induces a slight increase in the probability (relative to the other experiment groups) that the shown estimate is correct which does not materially affect the results (discussed further in Section 4).

[7] This was pre-registered, under the condition that the treatment effects in *Minimal* were minimal.

#0011924). All experiments were reviewed and approved by the IRB at the Norwegian School of Economics (NHH-IRB 39/22 and 44/22).

To improve data quality, participants were allowed to participate in the incentivised tasks only if they passed two attention checks and correctly answered a battery of comprehension questions. Participants are given two opportunities to pass the test. The pass rate in the *Religion* and *Caste* experiments is about 25%, and about 45% in the Prolific experiments.[8] Participants also respond to a series of questions including demographic characteristics, religiosity, attitudes towards caste-based affirmative action, exposure to people from other castes/religions, and beliefs about the abilities of different caste/religious groups on the experimental task in the *Caste* experiment. Details are presented in Appendix C.

In 2.2% of the decisions, participants' first-period decisions were the same as the source's decision. As the motive behind the second-period decision cannot be clearly attributed in such cases, these decisions are excluded from most of the analyses reported in this paper.[9]

2.4 Descriptive Patterns

This part of the paper presents aggregate patterns in participants' task behaviour in the experiments.

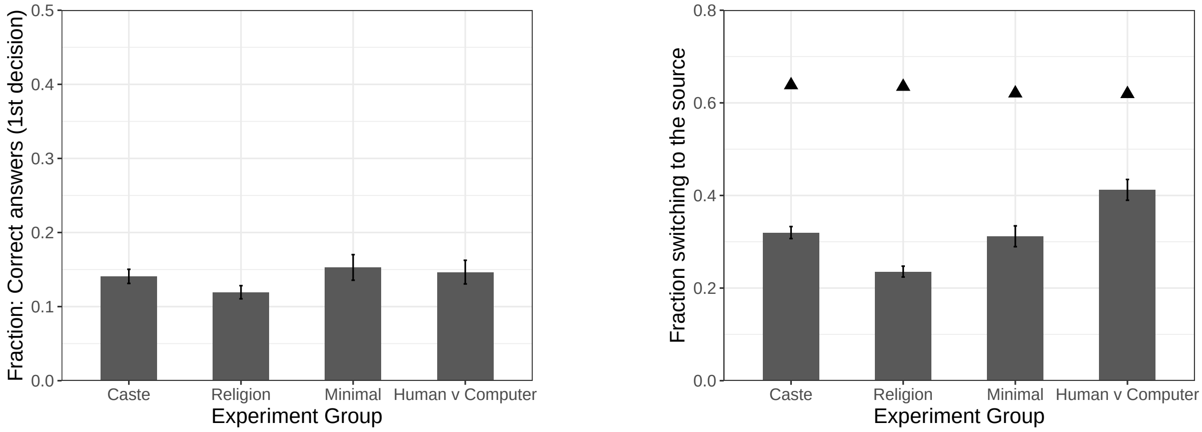
First-period performance. The bars in Panel (a) of Figure 2 show the fraction of first-period decisions that are objectively correct (incentive-yielding, within ±2 of y_T) in each experiment. The share of correct decisions ranges from $\approx$ 12% in the Religion experiment to $\approx$ 14 − 15% in the other experiments. Thus, while some participants make accurate decisions, most participants stand to improve their outcomes if they choose to switch in the second period. These patterns show that learning from others in this situation could be useful, fulfilling one of the design goals. The low success rates also indicate that most participants do not resort to simply using a spreadsheet or calculator to determine the correct answers.

Second-period behaviour. The bars in Panel (b) of Figure 2 show the fraction of participants who choose to learn by switching to the source in each experiment group. The triangles in the plot represent the fraction of participants who would improve their outcomes (i.e. earn more money) if they chose to switch. The figure shows that while a considerable fraction of participants switch (ranging from $\approx$ 24% in the Religion experiment to $\approx$ 43% in the Human vs. Computer experiment), a lot of participants leave money on the table by sticking with their own decision rather than switching to the source in the second period. Appendix Figure A.1 shows the distribution of the fraction of switching decisions in each experiment group. The figure shows that in all experiments, a small fraction of participants always switch, and in 3 of the 4 experiments, the modal participant never switches.

[8] The task comprehension rates, while low, are very similar to the comprehension rates reported in other research using the balls-and-urns task. For example, Enke and Graeber (2023) report a passing rate of 46% from a Prolific sample on a task that builds on the balls-and-urns paradigm.

[9] The results are unaffected when these observations are included, and when the exclusion window is expanded to within ±2.

Figure 2: Descriptive patterns: Task performance and switching rates.



(a) First period: Decision accuracy

(b) Second period: Switching

Notes. Panel (a) Bars indicate the fraction of decisions where participants make an incentive-yielding decision in the first period in an experiment group. (b): Bars indicate the fraction of participants who chose to switch to the source in the second period in an experiment group. Triangles indicate the fraction of decisions where switching to y_s would have been incentive-yielding. Error bars indicate 95% confidence intervals.

3 The Role Of Beliefs

This section presents experimental and survey evidence on the role of beliefs about the quality of information on learning when DMs and information sources share the same information.

Beliefs about the quality of information. Participants in the experiment receive a signal about the quality of the information provided by the source, which is the probability that the shown decision is correct. This is randomly assigned for each task and varies between 50% and 90%. Figure 3 shows the proportion of switch decisions at each level of the quality signal, separately for each of the experiments. As the quality of information increases from 50% to 90%, switching in the *Caste* and *Religion* experiments increases by ≈ 6% points and by ≈ 30% points in the *Minimal* and *Human vs. Computer* experiments.[10]

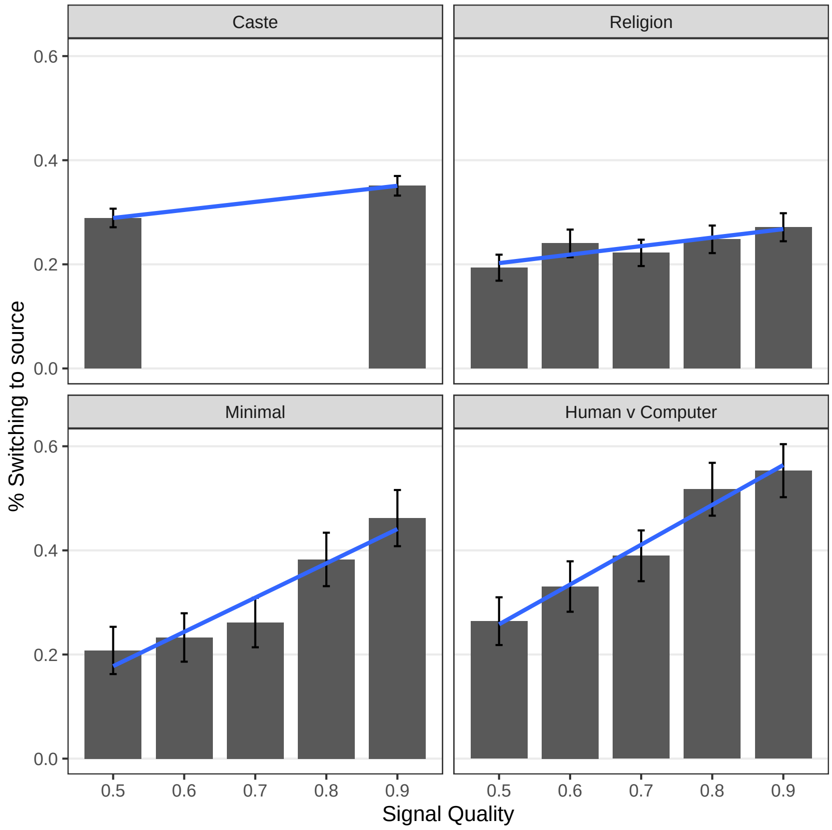
Table 2 presents results from regressions of the following specification for each experiment:

$$Switch_{i,n} = \beta_0 + \beta_1 Q_{i,n} + \beta_2 X_{i,n} + \gamma_i + T_i + \nu_n + \epsilon_i, \tag{1}$$

where *Switch* is an indicator variable for participant i in task n. $Switch = 1$ if the participant switches to the source in the second period and 0 otherwise. $Q_{i,n}$ is the level of the quality signal seen by participant i in task n. $X_{i,n}$ is a set of variables including the participant's confidence in their first period decision, indicator variables for whether their first guess was objectively correct and whether the source's shown decision was wrong, the difference between the participant's first decision and

[10]Appendix Figure A.2 reproduces this result controlling for the treatment condition to which participants are randomly assigned, and shows virtually the same patterns.

Figure 3: Effect of beliefs about information quality on learning



Notes. The graph shows the fraction of decisions where participants choose to switch to the source in each experiment, at each level of source quality Q. $Q \in \{50, 90\}$ in *Caste*, and $Q \in \{50, 60, 70, 80, 90\}$ in the other experiments. Error bars indicate 95% confidence intervals.

the source's decision, and the time spent while making the switching decision. T_i is an indicator for the treatment group that participants are randomly assigned to: $T_i = 1$ if the participant is assigned to see a source belonging to the General Caste (in *Caste*), is a Hindu (*Religion*), from their In-group (*Minimal*), or a Computer (*Human vs. Computer*). γ_i is a set of demographic controls (age, gender, tertiary education, and employment status). ν_n is a vector of task-specific controls: the order and a set of fixed effects for the task itself. Standard errors are heteroscedasticity-robust, clustered at the participant level. The coefficient of interest is β_1, which identifies the causal effect of an increase in the quality of information on switching. β_2 captures the partial correlations between switching and other variables of interest such as subjective confidence, conditional on signal quality.

The results show that the relationship between the quality signal and switching is statistically significant in all experiments – participants are more likely to switch when quality is higher. These results can be given a causal interpretation as the quality signal is exogenously assigned at the task level. Figure 3 also shows that switching is relatively low even when information quality is high – when the quality of the shown information is 90%, switching ranges from ≈ 36% to ≈ 53% in the different experiments. These patterns suggest that participants find switching costly, and are consistent with the findings from Weizsäcker (2010) and Conlon et al. (2022) which have documented that people do not learn from others even though they would benefit by doing so. The responsiveness to an increase in quality varies across samples and is more pronounced in experiments *Minimal* and *Human vs Computer*.

The balls-and-urns setup also enables an investigation of mechanisms that are the subject of active research on social learning and belief updating – confidence, objective performance, asymmetric updating, and attention.

Confidence and performance. Participants' reluctance to switch to the source may be partially driven by over-confidence. The median participant reports a confidence level of $\approx 85\%$ in the Caste and Religion experiments and $\approx 70\%$ in the Minimal and Human vs. Computer experiments that their first-period decision will earn the incentive. Participants behave in line with these beliefs: the estimated coefficient on confidence is negative and statistically significant in all four experiments, meaning that higher confidence in one's first period decision is linked with a lower rate of switching. Appendix Figure A.3 shows this visually – Panel (a) shows that more confident people make smaller errors (measured by the difference between their first-period decision and y_T), and Panel (b) shows that more confident people are less likely to switch in the second period. Participants who make accurate first-period decisions are also much less likely to switch than participants who fail to make an accurate decision. Thus, while the high levels of confidence are justifiable to some degree (as they are correlated with smaller errors), for most participants this results in lower earnings as only 13% of first-period decisions are accurate. Appendix Figure A.4 highlights this further, showing that participants switch more when information quality is higher than their confidence in their own decision. The consistency of these patterns across all of the experiments supports the conclusion that confidence is a powerful moderator of the switching decision.

Asymmetric learning and inferring quality from source decisions. Research on asymmetric belief updating has yielded differing results in different decision-making contexts (Benjamin, 2019; Barron, 2021). The balls-and-urns setup provides an opportunity to examine asymmetric updating in a situation with shared information and incentives. The estimated coefficients in Table 2 provide suggestive evidence of asymmetric updating for disconfirming signals – In 3 of the 4 experiments, participants switch more often when the difference between the source's decision y_S and their first period decision y_1 is larger. This suggests that participants may not be willing to abandon their decision for small adjustments, but are more likely to switch when the differences are larger. Another explanation for this pattern is that participants may be drawing inferences based on the estimate and the quality signal. For example, seeing a decision that is far away from their own first-period decision could induce a participant to switch differently than if seeing an estimate that is closer if they hold any beliefs about the correct posterior probability. The estimated coefficient on an indicator of whether the shown source is incorrect is negative, suggesting that participants may draw some information from the shown decisions independent of the quality signal. The results in Table 2 indicate that these channels may both be present.[11]

[11]Consequently, this is included as an additional control in all regressions although its omission does not affect any of the results as source quality is randomly assigned and the set of decisions linked to a task are the same for all participants.

Table 2: Regression analysis: The relationship between beliefs and switching.

	Dep. var.: Switch to source			
	Caste	Religion	Minimal	Computer
	(i)	(ii)	(iii)	(iv)
Quality	0.128***	0.125***	0.654***	0.665***
	(0.039)	(0.045)	(0.091)	(0.063)
Confidence	-0.003***	-0.002***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.001)	(0.000)
Guess 1 correct	-0.096***	-0.039**	-0.110***	-0.132***
	(0.021)	(0.019)	(0.029)	(0.021)
Source is wrong	-0.019	-0.036***	-0.049*	-0.058***
	(0.017)	(0.014)	(0.026)	(0.018)
Dist. to Source	0.000	0.204***	0.173**	0.141***
	(0.037)	(0.038)	(0.071)	(0.049)
Decision time (sec.)	0.002***	0.002***	0.003***	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.369***	0.281***	0.087	0.123
	(0.075)	(0.071)	(0.113)	(0.077)
Task controls	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓
Treatment FE	✓	✓	✓	✓
R^2	0.046	0.041	0.133	0.149
Dependent variable mean	0.317	0.234	0.317	0.370
Observations	4,754	4,963	1,576	3,364
Individuals	851	853	278	593

Notes. Estimates from OLS regressions where the dependent variable is whether a participant switches in the second period. The main variable of interest is *Quality*, which is the experimentally assigned quality signal. Other variables of interest are (i) participant's reported confidence on a given task (0-100), (ii) whether their first period guess was accurate (0/1), (iii) whether they saw an incorrect source (0/1), the difference between their decision and the source ($|y_1 - y_S|$), and (v) the time spent when deciding whether or not to switch. Controls are task and order fixed effects, demographic characteristics (the participant's age and dummies for employment, college education and gender), and controls for the identity treatment within each experiment. Observations with an extremely high decision time are excluded from these estimates. Robust standard errors, clustered at the participant level. $* p < 0.10, ** \ p < 0.05, *** \ p < 0.01$.

Attention and response times. The estimated coefficients on decision time in Table 2 show that participants who spend more time on the decision page are more likely to switch. Under the assumption that time spent deciding whether to switch is a reasonable proxy for attention, this correlation suggests that source decisions are not merely a substitute for the cognitively demanding process of computing the posterior probabilities, and that switching is not driven purely by a desire to complete the survey quickly.

Summary. These results show that (i) beliefs about the quality of information causally influence learning, (ii) learning rates are low, even when information quality is high, and (iii) the trade-off between information quality and subjective confidence is an important mechanism that moderates learning. At the same time, the results show that learning when both information sources and decision makers have the same information shares many of the systematic features documented in other decision-making contexts.

3.1 Beliefs and identity

In a separate survey ($N = 327$) conducted with Indian participants in July 2023, I elicited incentivised beliefs about the likelihood that a randomly chosen General Caste, Scheduled Caste, or Muslim individual would make an accurate first-period decision.[12] Most participants vastly overestimate the likelihood of success of all identity groups. High-status caste Hindus believe that the likelihood of success is 78.6% for fellow high-status Hindus, 71% for low-status Hindus, and 61.7% for Muslims, on average. These averages are very similar for low-status caste Hindus as well, suggesting that Hindus in our sample possess beliefs that high-status caste Hindus outperform low-status caste Hindus, and that Hindus (of all caste groups) outperform Muslims. These beliefs are much higher than the actual success rates of about $\approx 12-13\%$.

Figure 4 plots the cumulative distribution of these beliefs and shows that participants believe that people of different castes and religions differ in task performance. Panel (a) shows that participants believe that the G category outperforms the O category, and Panel (b) shows that participants believe that Hindus outperform Muslims.[13]

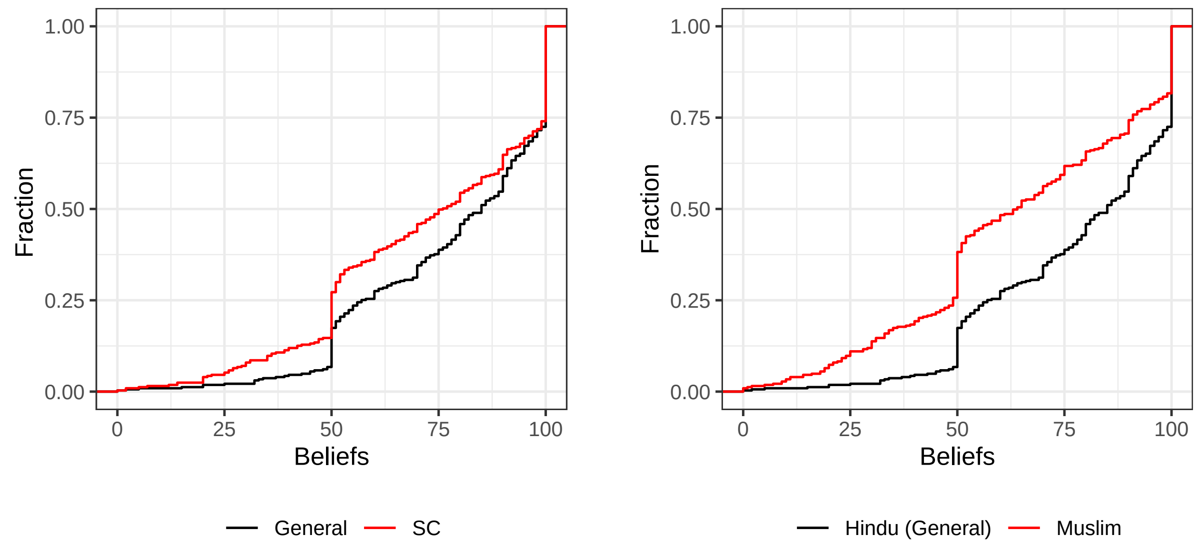
Beliefs without the quality signal. A supplementary experiment, *Caste – No Signal*, was conducted concurrently with experiment *Caste*. The only difference between these experiments was that participants were given a quality signal in the *Caste* experiment, but not in *Caste – No Signal*. In both experiments, participants were incentivised to provide their beliefs about the performance of the General and the SC caste groups after completing the main tasks (using the same question as in the beliefs survey discussed above).

Appendix Figure A.6 presents binned scatter plots of switching on participants' beliefs about the performance of a random individual from the caste group of the treatment to which they were assigned (either from the General category, or the Scheduled Castes category). The left panel of Appendix Figure A.6 shows that when the quality signal is not provided, switching increases with the participant's beliefs about success. The right panel of Appendix Figure A.6 shows that the correlation is virtually zero when the quality signal is provided. Strikingly, the increase in switching when beliefs about the success of the source's caste group moves from 50% to 90% is approximately the same as

[12] Participants in this survey only made first-period decisions on a single task, and thus were not exposed to sources.

[13] The p-values from two-sample Kolmogorov-Smirnov tests for the equality of distributions are < 0.01 in both cases. Appendix Figure A.5 shows this plot using data from the *Caste* experiment (in which beliefs were elicited after the decision tasks), which shows that the patterns are qualitatively very similar (KS test p-value < 0.01).

Figure 4: Beliefs about the performance of different groups



(a) Caste: High vs. low status

(b) Religion: Hindus vs. Muslims

Notes. Cumulative frequencies of participants' beliefs about the likelihood that a person from a particular caste category or religion makes an accurate first-period estimate in the task. Panel (a) – Red line: O category Hindu, Black line: G category Hindu. Panel (b) – Red line: Muslim, Black line: G category Hindu.

the difference induced by experimentally varying the signal quality by the same amount. However, as these beliefs are elicited after the main tasks, these results cannot be given a causal interpretation.

Together, this set of results provides evidence that participants over-estimate others' performance on the experimental task, and believe that performance differs by religion and caste. Thus, using identity to proxy for information quality could be misleading owing to misspecified performance beliefs. These analyses also highlight that failing to control beliefs about information quality when identity is salient confounds preference identification.

4 The Role Of Preferences

This section focuses on causally identifying preferences for the identity of the information source. The analysis begins with the average treatment effects in the four experiments, before moving on to in-group preferences in the *Caste* and *Religion* experiments, and concludes with an exploration of different sources of preference heterogeneity.

4.1 Preferences for the identity of information sources

I estimate the following specification using OLS regressions to causally identify preferences for the identity of information sources:

$$Switch_{i,n} = \beta_0 + \beta_1 T_i + Q_{i,n} + \gamma_i + \nu_n + \epsilon_i, \quad (2)$$

where *Switch* is an indicator variable for participant i in task n. *Switch* $= 1$ if the participant switches to the source in the second period and 0 otherwise. T_i is an indicator for the treatment group that participants are randomly assigned to: $T_i = 1$ if the participant is assigned to see a source belonging to the General Caste (in *Caste*), is a Hindu (*Religion*), from their In-group (*Minimal*), or a Computer (*Human vs. Computer*). $Q_{i,n}$ is the level of the quality signal seen by participant i in task n. γ_i is a set of demographic controls (age, gender, tertiary education, and employment status), and i's confidence in the accuracy of their first-period decision. ν_n is a vector of task-specific controls: the order, specific task, and whether the shown decision is correct. Standard errors are heteroscedasticity-robust, clustered at the participant level. The coefficient of interest is β_1, which identifies preferences for learning from one identity group over another.

Hypothesis 1. *The estimate of* $\beta_1 \neq 0$ *in each of the experiments, indicating that participants prefer to learn more from one of the two identity groups: (i) a General caste or SC/ST/OBC caste, (ii) a Hindu or a Muslim, (iii) experimentally assigned in-group or out-group, (iv) a Computer or an anonymous human.*

Table 3 presents the results from estimating this equation separately for each of the four main experiments. The estimated average treatment effect in the *Caste*, *Religion*, and *Minimal* experiments is statistically indistinguishable from zero, and the 95% confidence intervals are within 0.1 standard deviations. There is no evidence to support Hypothesis 1 (i)-(iii) that participants have preferences for whether the source is a *G* caste or *O* caste group individual in the *Caste* experiment, a Hindu or a Muslim in the *Religion* experiment, or their experimentally assigned in-group member or an out-group member in the *Minimal* experiment.

In contrast, the estimated treatment effect is positive and statistically significant for the *Human vs. Computer* experiment. The magnitude of the effect ($\approx 15\%$ points) is quantitatively large, about a third of the fraction switching in the *Human* condition. The results support Hypothesis 1 (iv) – participants prefer to switch more to decisions made by a computer or algorithm than by another human. This also supports "algorithmic appreciation", that humans prefer to learn from non-social sources when the quality of information is precisely known. The results from the *Human vs. Computer* experiment also demonstrate that the subtle identity manipulation – implemented through minor interface modifications – meaningfully shaped behaviour.

Robustness: Main results. The results are robust to the inclusion of different sets of controls and background characteristics. Detailed regression analyses – presented in Appendix Tables B.3–B.6 – show that the estimated average treatment effects are not sensitive to the choice of control variables. Columns (v) and (vi) of Appendix Table B.6 show that the results with and without including decisions from the *Minimal* experiment in the Human treatment for the *Human vs. Computer* experiment are virtually the same.

Table 3: Regression analysis: Preferences for the identity of information sources.

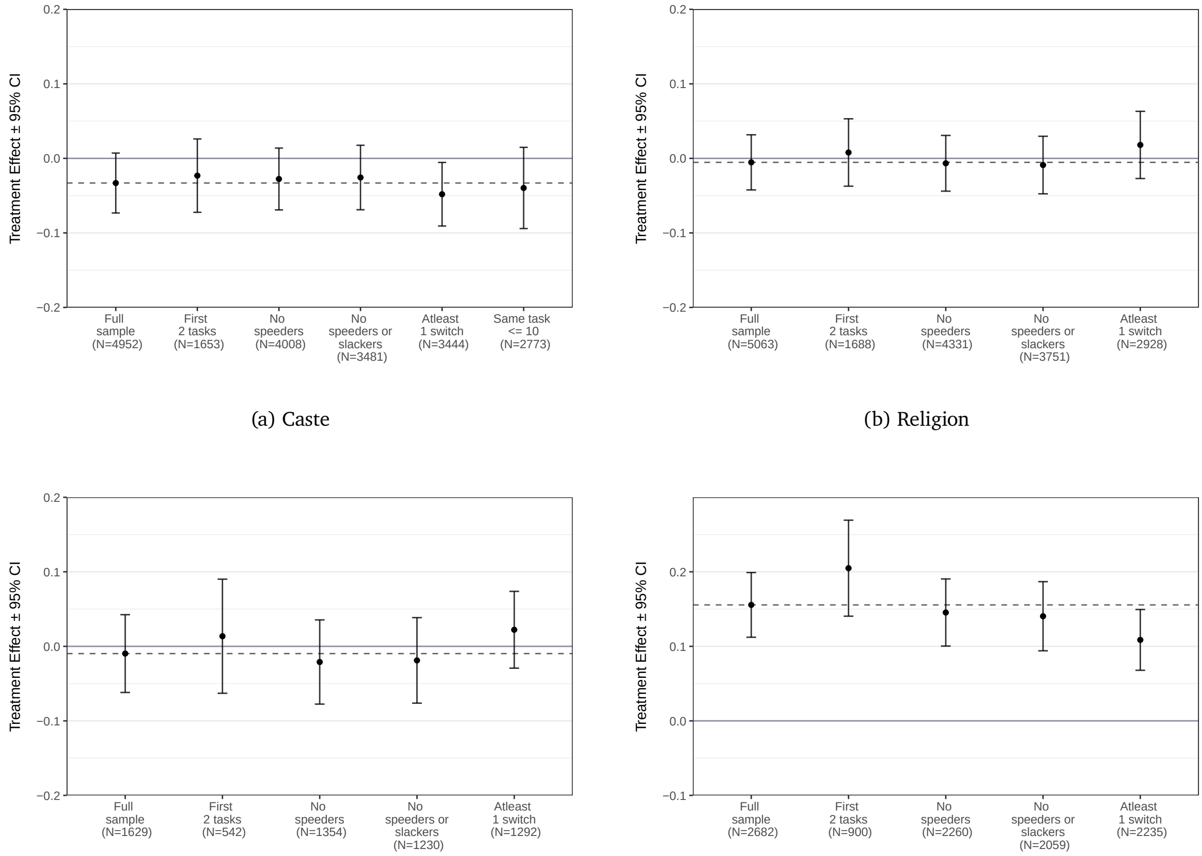
	Dep. var.: Switch to source			
	Caste	Religion	Minimal	Computer
	(i)	(ii)	(iii)	(iv)
General Caste Source	-0.033			
	(0.021)			
Hindu Source		-0.006		
		(0.019)		
In-group Source			-0.010	
			(0.027)	
Computer Source				0.156***
				(0.021)
General Caste	-0.015			
	(0.022)			
Hindu		-0.042*		
		(0.023)		
Quality	0.140***	0.127***	0.614***	0.642***
	(0.038)	(0.045)	(0.092)	(0.063)
Constant	0.431***	0.410***	0.241**	0.212***
	(0.072)	(0.072)	(0.101)	(0.071)
Task controls	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓
R^2	0.035	0.027	0.113	0.134
Dependent variable mean	0.320	0.236	0.312	0.365
Observations	4,952	5,063	1,629	3,473
Individuals	851	853	278	593

Notes. Average treatment effects in the different experiments, estimates from OLS regressions of Equation 2. The dependent variable is whether a participant switches in the second period. Controls are task and order fixed effects, whether the source is correct, and demographic characteristics (the participant's age and dummies for employment, college education and gender). Robust standard errors, clustered at the participant level. $* \, p < 0.10, ** \; p < 0.05, *** \, p < 0.01$.

Figure 5 shows the results from a battery of additional robustness tests examining whether the results are driven by a lack of comprehension of the task, confusion, or inattention – a plausible concern given the cognitively demanding nature of the balls-and-urns task. These tests use features of participants' decisions that could represent random decision making, cognitive fatigue, or a lack of understanding of the task and incentive structure. The panels show estimates of the average treatment effect for each experiment under different sample restrictions using the specification in Equation 2.

First, participants may be more motivated, energetic, or likely to remember instructions at the beginning of the experiment. Restricting the sample to the first two (of six) tasks does not affect the estimated treatment effect relative to the full sample. Next, completing tasks very quickly or very slowly may reflect inattention or distraction. The third and fourth estimates in each panel exclude decisions that are in the top and bottom 10% of time taken while making the switching decision. Excluding such decisions has virtually no effect on the estimates. Next, participants who never switch

Figure 5: Robustness tests



(a) Caste

(b) Religion

(c) Minimal Identity

(d) Human vs. Computer

Notes. Each point is the estimated coefficient on the treatment dummy variable from OLS regressions of Equation 2 under different sample restrictions. Error bars indicate 95% confidence intervals. Standard errors are clustered at the participant level. Controls are: A set of demographic controls, task and order fixed effects, confidence, whether the source was correct, and the quality of the information source. The dashed lines are a reference for the point estimate of the average treatment effect using the main estimation sample.

may arguably do so because they just want to get through the tasks quickly. Restricting the sample to participants who switch at least once, the estimates are very similar in three of the four experiments. The estimate is statistically significant (and negative) at the 5% level in the *Caste* experiment.

Finally, experiment *Caste* included a design feature to specifically examine participant inattention. The last of the six tasks faced by a participant was the same as one of the first 3 tasks. The correlation between participants' responses on the identical tasks is 0.62, which is quite strong. Further, the estimated treatment effect on a restricted subsample of participants whose decisions on identical tasks differ by less than 10 – an indicator of high consistency – is very similar to the estimate using the full sample.

Taken together, the consistency of the results across all of these robustness tests support the main results that participants do not have a preference for the identity of the information source, and that they have a preference for information from non-social sources over information from another human.

Do preferences emerge when beliefs are not controlled? The results on beliefs in Sections 2.4 and 3 show that in the absence of a quality signal, participants may rely on underlying beliefs about the performance of different caste groups (which are higher for the General caste group). Does learning because of the caste identity of the source differ within experiment *Caste – No Signal*? Appendix Table B.7 shows the results from estimating the specification in Equation 2 for this experiment. The results show that the estimated treatment effect is not statistically significant – participants do not appear to switch differently whether they see a General Caste or a SC/ST/OBC source.

To place these results in perspective, note that in the *Caste* experiment a 40% (exogenous) increase in source quality leads to a $\approx 6\%$ increase in switching. In comparison, participants' beliefs about the differences in ability between the two caste groups is much smaller, about $5-10\%$ on average. Thus, in the absence of preferences for the caste identity of information sources, the *No Signal* experiment (with a sample size of 295 participants) is underpowered to detect differences in switching because of the relatively small difference in beliefs about the performance of different groups.

4.2 In-group preferences and overlapping identities

A person's identity *relative* to the source of information could influence switching in the experiment, given previous research showing that identities affect decision-making through in-group favouritism and out-group parochialism (Shayo, 2020; Charness and Chen, 2020). While the *Minimal* experiment finds no evidence of in-group preferences with experimentally assigned identities, the *Religion* and *Caste* experiments allow for further investigation with naturally-occurring identities. Participants in these experiments self-report their caste group and religion, enabling an analysis of whether they prefer information from an in-group source (same caste or religion) over an out-group source (different caste or religion).

I estimate the following specification using OLS regressions to identify the existence of in-group preferences for the identity of the information source:

$$Switch_{i,n} = \beta_0 + \beta_1 T_i + \beta_2 I_i + \beta_3 T_i \times I_i + Q_{i,n} + \gamma_i + \mu_n + \epsilon_i, \tag{3}$$

where I_i is an indicator variable of the participant's identity: In experiment *Caste*, $I_i = 1$ if the participant belongs to the general category (G) and $I_i = 0$ if the participant belongs to the SC/ST/OBC categories. In experiment *Religion*, $I_i = 1$ if the participant is Hindu and $I_i = 0$ if the participant is Muslim. The sample for the *Religion* experiment is restricted to Hindu and Muslim participants to allow for the construction of clear in-groups.

The coefficient β_1 is the treatment effect of seeing a $T_i = 1$ source relative to a $T_i = 0$ source for $I_i = 0$ participants. $\beta_1 + \beta_3$ is the treatment effect of seeing a $T_i = 1$ source relative to a $T_i = 0$ source for $I_i = 1$ group participants. The interaction coefficient β_3 is the difference in the treatment effects between $I_i = 0$ and $I_i = 1$ identity group participants.

Hypothesis 2. *The estimate of $\beta_3 \neq 0$, indicating that participants learn differently from an in-group source than from an out-group source within the Religion and Caste experiments.*

Columns (i)-(ii) of Table 4 present the results from estimating equation 3 in the *Caste* and *Religion* experiment samples. The estimated coefficients on the interaction terms are not (individually or jointly) statistically significant. The results are robust to the inclusion of task and demographic controls, and to controlling for participant's stated confidence. Appendix Tables B.8 and B.9 present the results from same robustness tests and show that the results are largely consistent. Further, there is no evidence that participants belonging to different castes or religions switch differently, nor are the treatment effects for any of the subgroups statistically significant. Taken together, there is no evidence that participants belonging to different castes or religions have in-group preferences for information from their caste or religion in-groups.

Overlapping identities. The analysis so far has focused on a single dimension of identity. However, in naturally occurring contexts, people possess multiple identities which may affect learning in different ways. In the Indian context, people have both religious and caste identities. The importance of caste identity and the associated behavioural norms and prescriptions vary between caste groups and across religions. The *Religion* experiment includes participants of different religions who also hold caste identities, which makes it possible to investigate how preferences for the religious identity of sources vary across different overlaps of religious and caste identities.

I estimate the following specification to study the effect of overlapping religious and caste identities in the *Religion* experiment:

$$Switch_{i,n} = \beta_0 + \beta_1 T_i + \beta_2 C_i + \beta_3 T_i \times C_i + Q_{i,n} + \gamma_i + \mu_n + \epsilon_i, \tag{4}$$

where C_i is the caste identity of participant i. $C_i = 0$ if the participant belongs to the SC/ST/OBC caste category, and $C_i = 1$ if the participant belongs to the general caste category.

Columns (iii)-(iv) of Table 4 present estimates of Equation 4 using sub-samples of Hindu and Muslim participants within experiment *Religion*. Column (iii) shows that Hindu participants belonging to the SC/ST/OBC castes switch less when they see a Hindu source than when they see a Muslim source. The interaction term (Hindu Source x General Caste) is positive and statistically significant, indicating that (Hindu) General caste participants switch more to Hindu sources than Muslim sources when compared to SC/ST/OBC participants. Turning to the Muslim sub-sample (Column (iv)), there is no evidence that switching to Hindu or Muslim sources differs by the (Muslim) participant's caste group.

These results provide only weak evidence that preferences for the identity of information sources may exist for (or between) specific identity groups. The analysis also highlights a key challenge for research with naturally-occurring identities, that individuals belong to many identity groups concurrently which may have complex interactions. These results also suggest that demand effects, if any, are minimal as they would cause participants from different religious or caste groups to act in opposite directions for the same source identity.

Table 4: Regression analysis: Heterogeneous effects by caste category

	Dep. var.: Switch to source			
	In-group Preferences Experiment		Overlapping Identities Participant Religion	
	Caste	Religion	Hindu	Muslim
	(i)	(ii)	(iii)	(iv)
General Caste Source	-0.013			
	(0.035)			
General Caste Source × General Caste	-0.032			
	(0.043)			
Hindu Source		0.079	-0.105***	0.073
		(0.065)	(0.038)	(0.091)
Hindu Source × Hindu		-0.093		
		(0.068)		
Hindu Source × General Caste			0.141***	-0.005
			(0.046)	(0.128)
General Caste	0.000		-0.102***	0.009
	(0.031)		(0.035)	(0.077)
Hindu		0.003		
		(0.041)		
Quality	0.056***	0.091*	0.100**	0.054
	(0.015)	(0.049)	(0.051)	(0.164)
Constant	0.490***	0.372***	0.420***	0.321
	(0.068)	(0.085)	(0.085)	(0.240)
Task controls	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓
R^2	0.035	0.024	0.029	0.067
Dependent variable mean	0.320	0.230	0.225	0.270
Observations	4,952	4,348	3,841	507
Individuals	851	733	647	86

Notes. Columns (i)–(ii): Estimates of Equation 3 for in-group preferences in the *Caste* and *Religion* experiments. Columns (iii)–(iv): Estimates of Equation 4 for in-group preferences with overlapping identities using sub-samples of Hindu and Muslim participants in the *Religion* experiment. The dependent variable is 1 if the participant switches to the source's decision y_s. Demographic controls are the participant's age and dummies for employment, college education and gender. Task controls include task and order effects, and an indicator for whether the source is correct. Standard errors are clustered at the participant level. $* \, p < 0.10, ** \, p < 0.05, *** p < 0.01$.

4.3 Mechanisms and heterogeneity

Behavioural mechanisms

I begin by following a similar approach as in Section 3 by examining whether several mechanisms which affect learning and belief formation vary by the identity of the information source. The mechanisms include experimentally controlled parameters and endogenous responses: (i) experimental variation in the quality of the information (whether responsiveness to quality varies by source identity), (ii) experimental variation in whether the source is correct or incorrect (whether participants discern incorrect sources differently based on identity), (iii) the participant's confidence in their first-period decision (whether the quality-confidence trade-off varies by identity), (iv) the accuracy of the participant's first-period decision (whether high or low ability participants react differently to source identity), (v) the difference between the first-period decision and the source's decision ($|y_1 - y_S|$, whether responses to sources whose estimates confirm or disconfirm participants' estimates differ by identity), and (vi) the time spent when making the switching decision (whether more deliberative decisions differ by source identity). Conceptually, evidence that these mechanisms affect behaviour based on the source's identity would support the existence of identity preferences that emerge based on contextual features when making these decisions.

I estimate the following specification using OLS to explore the interaction between a mechanism and the source's identity:

$$Switch_{i,n} = \beta_0 + \beta_1 T_i + \beta_2 M_{i,n} + \beta_3 T_i \times M_{i,n} + \gamma_i + \mu_n + \epsilon_i, \tag{5}$$

where $M_{i,n}$ denotes whether participant i encounters mechanism M in task n. The coefficient on the interaction term β_3 can be interpreted as the difference in switching associated with the mechanism because of a difference in the source's identity. Appendix Tables B.10, B.11, B.12, and B.13 present the results from estimating this specification for each of the 6 mechanisms within each experiment.

Across experiments, the results provide very little evidence that these mechanisms interact with identity, with only 3 of the 24 estimated coefficients being statistically significant at the 5% level. This stands in contrast to the strong and systematic impact of these mechanisms, discussed in Section 3. There is no evidence of heterogeneous treatment effects by the quality of the information provided by the source. Appendix Figure A.7 shows that even in the *Human vs. Computer* experiment, where there are large and robust average effects, the treatment difference is very similar at all levels of source quality.[14]

In the *Religion* experiment, the results in Column (i) of Table B.11 show that participants may respond more to an increase in the quality of a Hindu source than a Muslim source. An analysis of this effect (not reported) shows that the difference in responsiveness to the quality of information appears to be driven by General category Hindus. SC/ST/OBC Hindus and Muslims of all caste groups do not appear to be differentially sensitive based on the source's religion.

In the *Human vs. Computer* experiment, the results in columns (v) and (vi) of Table B.13 suggest that participants may switch less to a computer than to another human when (i) the distance between the source decision and the participant's first-period decision increases, and (ii) when they spend more time on the switching decision.

Overall, the lack of heterogeneity because of source identity suggests that preferences for the identity of information sources do not play a role when people decide whether to learn.

[14]An interesting consequence of this behaviour is that in the *Computer* condition, participants actually switch more than they should when information quality is low (i.e. only a 50% chance of being correct).

Strength of caste identity

How strongly people identify with their caste identity may be related to their preferences for the (caste) identity of information sources. To explore this, I elicited two measures strongly linked with the strength of an individual's caste identity in experiment *Caste*: (i) the caste and religious composition of participants' friend networks, and (ii) their support for caste-based affirmative action (commonly known as "reservations", an important topic in the Indian context). Appendix Figure A.8 shows the estimated coefficients of the difference in the average treatment effect between people who have fewer friends from other castes, fewer friends from other religions, or are against caste-based affirmative action policies, relative to those who have more friends from other castes/religions or who support affirmative action policies.[15] None of the estimated coefficients are statistically significant, meaning that there is no evidence of heterogeneous preferences for the identity of information sources because of differences in the strength of caste identity.

Taken together, there is very little evidence for preference heterogeneity along various characteristics, which supports the main result that preferences for the identity of information sources do not influence whether people learn.

5 Conclusion

This paper studies how the identity of an information source affects whether people choose to learn, using online experiments that study both naturally occurring and experimentally assigned identities. The results show that while beliefs about information quality influence learning significantly, preferences for the identity of the information source do not play a role when the source is human. The results from the *Human vs. Computer* experiment indicate that people prefer to learn from a non-social source than from another human, and present an interesting research question to identify the mechanisms underlying this effect both theoretically and empirically.

The experimental design examines a novel form of learning where a rational benchmark predicts that no learning will occur. It provides evidence of an under-studied source of learning in situations where stakeholders and sources have access to the same information but may differ in how they interpret it. This channel is also likely to influence learning in situations where both private and public information are available. Characterising how learning is motivated by a quest for new information and for better interpretations of the same information presents an exciting research agenda, closely related to the growing literature on the role of narratives in economic decision-making (Shiller, 2017; Graeber et al., 2022; Barron and Fries, 2023).

The experiment developed in this paper is deliberately abstract, which helps eliminate the numerous channels through which identity or contextual factors could influence learning. While this approach enables an analysis of the different roles of preferences and beliefs, it does so in a stylised setting where there is no possibility of prolonged association with out-group members. On the other hand, the experiment design shares core elements with many everyday decisions – we read newspa-

[15] The estimated specification is $Switch_{i,n} = \beta_0 + \beta_1 In_i + \beta_2 E_i + \beta_3 In_i \times E_i + Q_{i,n} + \gamma_i + \mu_n + \epsilon_i$, where $In_i = 1$ if the participant and the source are in the same caste group and E_i is an indicator variable for each mechanism.

per articles, read online reviews, or watch news on the TV which may be delivered by people from different social groups. Future research can adapt the experiment introduced in this paper to study how identity influences social learning in contextually richer decision-making environments, involving additional important elements such as repeated interactions, ego-relevant outcomes, or multiple identities.

These results carry important implications for policy-makers and organisations. Although the results show that preferences for identity do not matter, this holds only when the quality of information is both salient and precise. As providing an unambiguously clear signal is difficult in most real-world situations, policy-makers would benefit from finding ways to minimise the role of messenger identity when designing information delivery campaigns and outreach efforts. Emphasising information quality and delivering information through non-social channels may yield better results in settings such as the take-up of social welfare programs, adoption of better health practices, or engineering changes in harmful social norms.

References

Afrouzi, Hassan, Carolina Arteaga, and Emily Weisburst, "Is It the Message or the Messenger? Examining Movement in Immigration Beliefs," *Journal of Political Economy Microeconomics*, May 2024, *2* (2), 244–297.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *Quarterly Journal of Economics*, aug 2000, *115* (3), 715–753.

Arrow, Kenneth J, "The Theory of Discrimination, Discrimination in Labor Markets, Ashenfelter, O. and A. Rees eds., 3-33," 1973.

Asher, Sam, Paul Novosad, and Charlie Rafkin, "Intergenerational Mobility in India: New Measures and Estimates across Time and Social Groups," *American Economic Journal: Applied Economics*, April 2024, *16* (2), 66–98.

Atkin, David, Eve Colson-Sihra, and Moses Shayo, "How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption," *Journal of Political Economy*, apr 2021, *129* (4), 1193–1251.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune, "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, may 2013, *5* (2), 33–72.

Barron, Kai, "Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains?," *Experimental Economics*, March 2021, *24* (1), 31–58.

— **and Tilman Fries**, "Narrative Persuasion," *SSRN Electronic Journal*, 2023.

Bauer, Kevin, Yan Chen, Florian Hett, and Michael Kosfeld, "Group Identity and Belief Formation: A Decomposition of Political Polarization," *SSRN Electronic Journal*, 2023.

Becker, Gary S., "The economics of discrimination," *University of Chicago Press*, 1957.

Benjamin, Daniel J., "Errors in probabilistic reasoning and judgment biases," in "Handbook of Behavioral Economics - Foundations and Applications 2," Elsevier, 2019, pp. 69–186.

BenYishay, Ariel and A Mushfiq Mobarak, "Social Learning and Incentives for Experimentation and Communication," *The Review of Economic Studies*, jul 2018, *86* (3), 976–1009.

Berger, Sebastian, Christoph Feldhaus, and Axel Ockenfels, "A shared identity promotes herding in an information cascade game," *Journal of the Economic Science Association*, may 2018, *4* (1), 63–72.

Bertrand, Marianne and Esther Duflo, "Field experiments on discrimination," *Handbook of economic field experiments*, 2017, *1*, 309–393.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope, “Inaccurate Statistical Discrimination: An Identification Problem,” *Review of Economics and Statistics*, sep 2023, pp. 1–45.

Boyd, Robert, Peter J. Richerson, and Joseph Henrich, “The cultural niche: Why social learning is essential for human adaptation,” *Proceedings of the National Academy of Sciences*, jun 2011, *108* (supplement_2), 10918–10925.

Cassan, G, D Keniston, and T Kleineberg, “A division of laborers: Identity and efficiency in India,” Technical Report feb 2019.

Census of India, “Census of India 2011 Provisional Population Totals.,” *New Delhi: Office of the Registrar General and Census Commissioner*, 2011.

Chakravarty, Surajeet, Miguel A. Fonseca, Sudeep Ghosh, and Sugata Marjit, “Religious fragmentation, social identity and cooperation: Evidence from an artefactual field experiment in India,” *European Economic Review*, November 2016, *90*, 265–279.

Charness, Gary and Yan Chen, “Social Identity, Group Behavior, and Teams,” *Annual Review of Economics*, aug 2020, *12* (1), 691–713.

Chen, Daniel L., Martin Schonger, and Chris Wickens, “oTree—An open-source platform for laboratory, online, and field experiments,” *Journal of Behavioral and Experimental Finance*, mar 2016, *9*, 88–97.

Chen, Yan and Sherry Xin Li, “Group identity and social preferences,” *American Economic Review*, 2009, *99* (1), 431–57.

_ , _ , Tracy Xiao Liu, and Margaret Shih, “Which hat to wear? Impact of natural identities on coordination and cooperation,” *Games and Economic Behavior*, 2014, *84*, 58–86.

Conlon, John J, Malavika Mani, Gautam Rao, Matthew Ridley, and Frank Schilbach, “Failing to Learn from Others,” 2022.

Conlon, John, Malavika Mani, Gautam Rao, Matthew Ridley, and Frank Schilbach, “Learning in the Household,” Technical Report may 2021.

Danz, David, Lise Vesterlund, and Alistair J. Wilson, “Belief Elicitation and Behavioral Incentive Compatibility,” *American Economic Review*, September 2022, *112* (9), 2851–2883.

Dekel, Inbal and Moses Shayo, “Follow the Crowd: But Who Follows, Who Counteracts, and Which Crowd?,” *Who Counteracts, and Which Crowd*, 2023.

Dhami, Sanjit, Mengxing Wei, and Pavan Mamidi, “Religious identity, trust, reciprocity, and prosociality: Theory and evidence,” *Journal of Development Economics*, January 2024, *166*, 103192.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey, “Algorithm aversion: People erroneously avoid algorithms after seeing them err.,” *Journal of Experimental Psychology: General*, 2015, *144* (1), 114–126.

— , — , **and** — , “Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them,” *Management Science*, mar 2018, *64* (3), 1155–1170.

Dolan, P., M. Hallsworth, D. Halpern, D. King, R. Metcalfe, and I. Vlaev, “Influencing behaviour: The mindspace way,” *Journal of Economic Psychology*, February 2012, *33* (1), 264–277.

Enke, Benjamin and Thomas Graeber, “Cognitive Uncertainty,” *The Quarterly Journal of Economics*, may 2023, *138* (4), 2021–2067.

Fershtman, C. and U. Gneezy, “Discrimination in a Segmented Society: An Experimental Approach,” *The Quarterly Journal of Economics*, February 2001, *116* (1), 351–377.

Garcia-Hombrados, Jorge, Marcel Jansen, Ángel Martínez, Berkay Özcan, Pedro Rey-Biel, and Antonio Roldán-Monés, “Ideological alignment and evidence-based policy adoption,” Technical Report, IZA Discussion Papers 2024.

Ghosh, Arkadev, “Religious Divisions and Production Technology: Experimental Evidence from India,” *SSRN Electronic Journal*, 2022.

Goeree, Jacob K. and Leeat Yariv, “Conformity in the lab,” *Journal of the Economic Science Association*, feb 2015, *1* (1), 15–28.

Graeber, Thomas, Florian Zimmermann, and Christopher Roth, “Stories, Statistics, and Memory,” *SSRN Electronic Journal*, 2022.

Guryan, Jonathan and Kerwin Kofi Charles, “Taste-Based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots,” *The Economic Journal*, nov 2013, *123* (572), F417–F432.

Henrich, Joseph, *The Secret of Our Success*, Princeton University Press, dec 2016.

Hoff, Karla, Mayuresh Kshetramade, and Ernst Fehr, “Caste and punishment: The legacy of caste culture in norm enforcement,” *The economic journal*, 2011, *121* (556), F449–F475.

Hou, Yoyo Tsung-Yu and Malte F. Jung, “Who is the Expert? Reconciling Algorithm Aversion and Algorithm Appreciation in AI-Supported Decision Making,” *Proceedings of the ACM on Human-Computer Interaction*, oct 2021, *5* (CSCW2), 1–25.

Iyer, Sriya, “The New Economics of Religion,” *Journal of Economic Literature*, June 2016, *54* (2), 395–441.

— **and Anand Shrivastava**, “Religious riots and electoral politics in India,” *Journal of Development Economics*, March 2018, *131*, 104–122.

Jaffrelot, Christophe, *Modi’s India: Hindu nationalism and the rise of ethnic democracy*, Princeton University Press, 2021.

Langer, Markus, Tim Hunsicker, Tina Feldkamp, Cornelius J. König, and Nina Grgić-Hlača, ""Look! It's a Computer Program! It's an Algorithm! It's AI!": Does Terminology Affect Human Perceptions and Evaluations of Algorithmic Decision-Making Systems?," in "CHI Conference on Human Factors in Computing Systems" CHI '22 ACM April 2022, pp. 1–28.

Levy, Gilat and Ronny Razin, "Echo Chambers and Their Effects on Economic and Political Outcomes," *Annual Review of Economics*, aug 2019, *11* (1), 303–328.

Logg, Jennifer M, Julia A Minson, and Don A Moore, "Algorithm appreciation: People prefer algorithmic to human judgment," *Organizational Behavior and Human Decision Processes*, 2019, *151*, 90–103.

Maclean, Johanna Catherine, John Buckell, and Joachim Marti, *Information Source and Cigarettes: Experimental Evidence on the Messenger Effect* March 2019.

Mobius, Markus and Tanya Rosenblat, "Social Learning in Economics," *Annual Review of Economics*, aug 2014, *6* (1), 827–847.

Mosse, David, "The Modernity of Caste and the Market Economy," *Modern Asian Studies*, oct 2019, *54* (4), 1225–1271.

Munshi, Kaivan, "Caste and the Indian Economy," *Journal of Economic Literature*, dec 2019, *57* (4), 781–834.

Oh, Suanna, "Does Identity Affect Labor Supply?," *American Economic Review*, aug 2023, *113* (8), 2055–2083.

Parys, Jessica Van and Elliott Ash, "Sequential decision-making with group identity," *Journal of Economic Psychology*, dec 2018, *69*, 1–18.

Pew Research Center, "Religion in India: tolerance and segregation," in "Pew Forum" 2021.

Phelps, Edmund S, "The statistical theory of racism and sexism," *The american economic review*, 1972, *62* (4), 659–661.

Robbett, Andrea, Lily Colón, and Peter Hans Matthews, "Partisan political beliefs and social learning," *Journal of Public Economics*, apr 2023, *220*, 104834.

Sankaran, Sindhuja, Maciek Sekerdej, and Ulrich von Hecker, "The Role of Indian Caste Identity and Caste Inconsistent Norms on Status Representation," *Frontiers in Psychology*, mar 2017, *8*.

Shayo, Moses, "Social Identity and Economic Policy," *Annual Review of Economics*, aug 2020, *12* (1), 355–389.

Shiller, Robert J., "Narrative Economics," *American Economic Review*, apr 2017, *107* (4), 967–1004.

Siddique, Zahra, "Evidence on Caste Based Discrimination," *Labour Economics*, dec 2011, *18*, S146–S159.

Tajfel, Henri and John C Turner, "Intergroup behavior," *Introducing social psychology*, 1978, *401*, 466.

__ , **M. G. Billig, R. P. Bundy, and Claude Flament**, "Social categorization and intergroup behaviour," *European Journal of Social Psychology*, 1971, *1* (2), 149–178.

Thorat, Sukhadeo and Paul Attewell, "The legacy of social exclusion: A correspondence study of job discrimination in India," *Economic and political weekly*, 2007, pp. 4141–4145.

Triandis, Harry C and Michele J Gelfand, "Converging measurement of horizontal and vertical individualism and collectivism.," *Journal of personality and social psychology*, 1998, *74* (1), 118.

Wabitsch, Alena, "The Messenger Matters," Technical Report 2024.

Weizsäcker, Georg, "Do We Follow Others when We Should? A Simple Test of Rational Expectations," *American Economic Review*, dec 2010, *100* (5), 2340–2360.

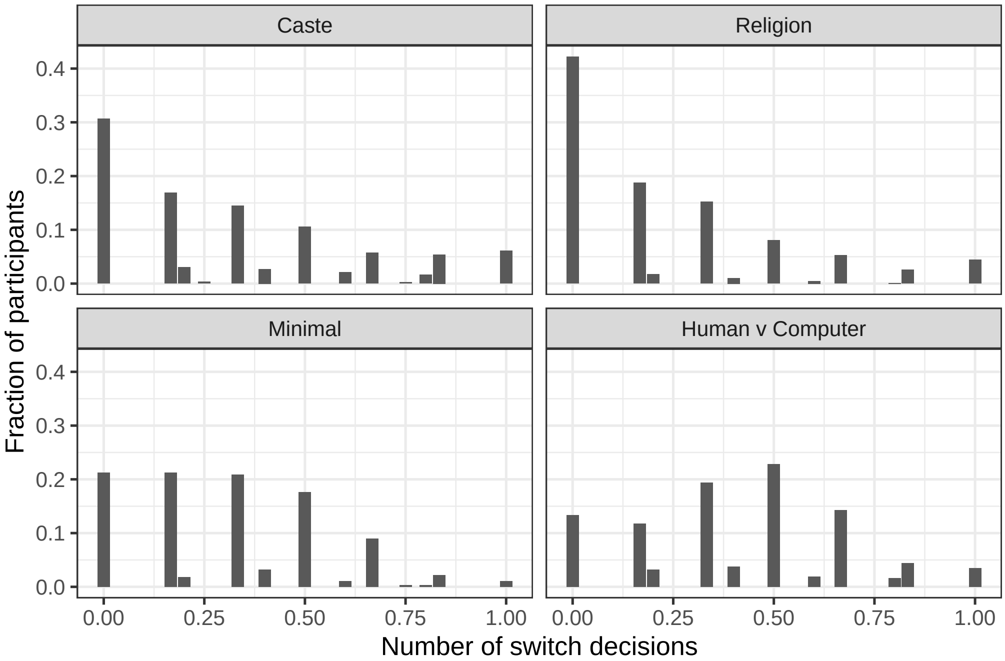
Zhang, Yunhao and David G. Rand, "Sincere or motivated? Partisan bias in advice-taking," *Judgment and Decision Making*, 2023, *18.*

Zou, Wenbo and Xue Xu, "Ingroup bias in a social learning experiment," *Experimental Economics*, dec 2022, *26* (1), 27–54.

APPENDIX

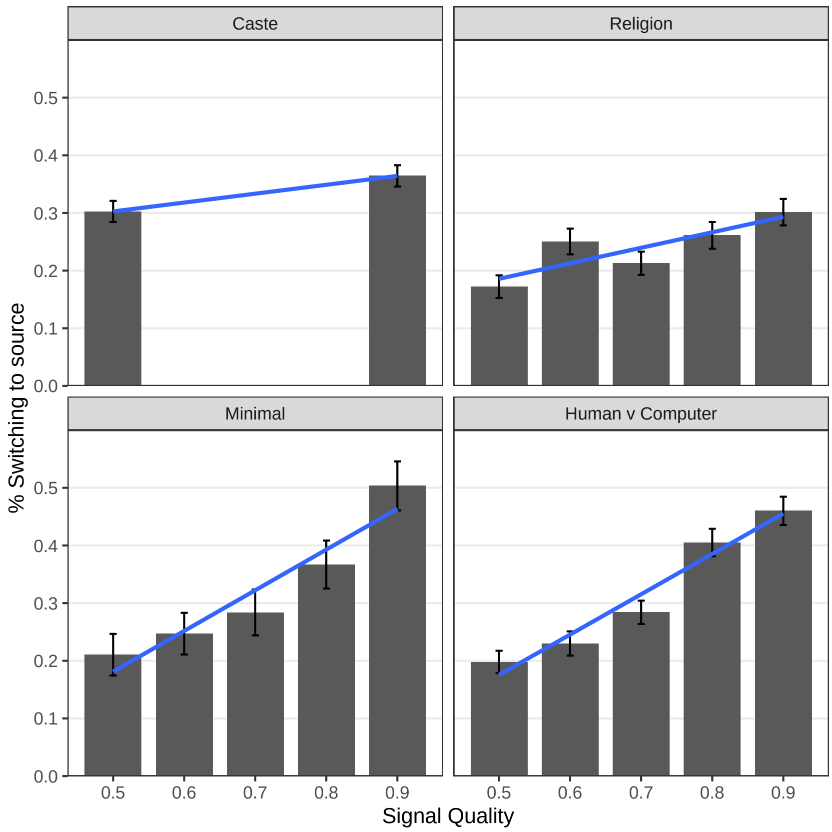
A Additional Figures

Figure A.1: Distribution of switching decisions – participant level



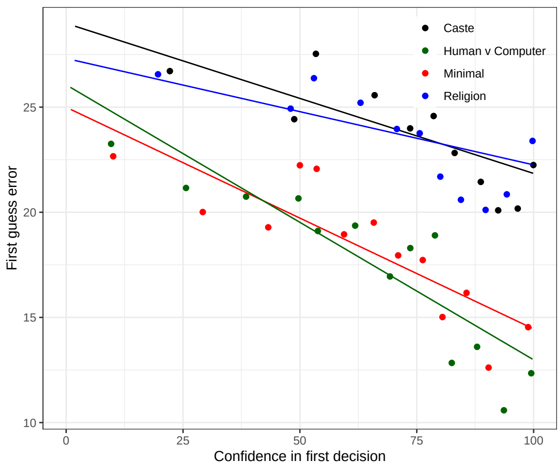
Notes. The graph shows the participant-wise distribution of the fraction of second-period decisions in which they choose to switch to the source. Note that in some cases the fraction is not a multiple of 1/6 as some decisions are excluded.

Figure A.2: Effect of beliefs about information quality on learning (controlling for differences between treatments)

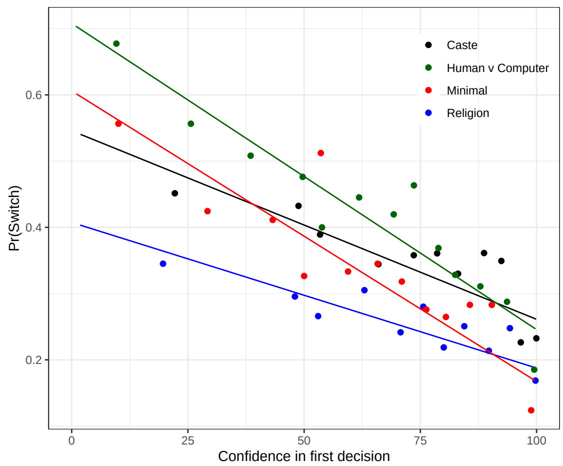


Notes. The graph shows the fraction of decisions where participants choose to switch to the source in each experiment, at each level of source quality Q, controlling for the treatment to which participants are assigned to. $Q \in \{50, 90\}$ in *Caste*, and $Q \in \{50, 60, 70, 80, 90\}$ in the other experiments. Error bars indicate 95% confidence intervals.

Figure A.3: Confidence: Task performance and switching



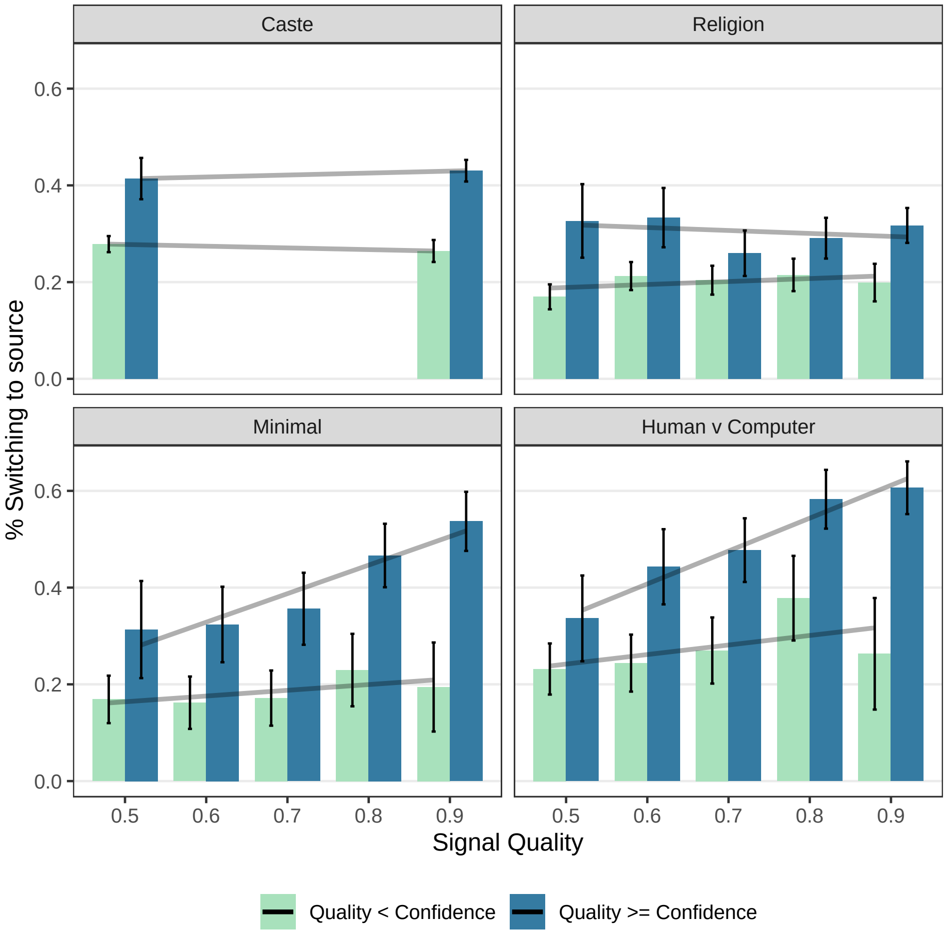
(a) First period: Accuracy



(b) Second period: Switching

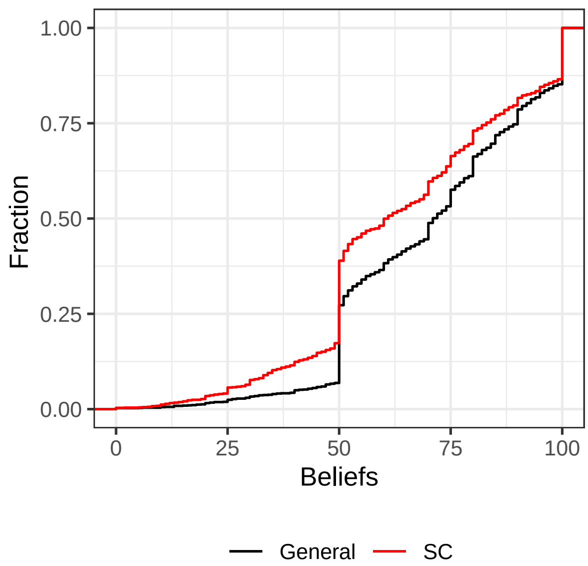
Notes. Binned scatter plots showing the relationship between the participants' stated confidence in the accuracy of their first period guess and (a): the error in the first period decision, and (b): the likelihood of switching to the source's estimate. Results are shown pooling treatment conditions in each experiment.

Figure A.4: Confidence vs. quality: Switching patterns



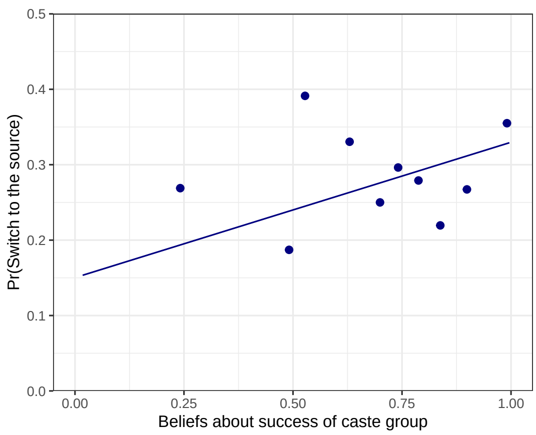
Notes. The graph shows the fraction of decisions where participants choose to switch to the source by the quality of the source. The light coloured bars are decisions when source quality is lower than participants' confidence in their first-period decision, dark coloured bars are decisions where quality is higher than confidence. Error bars indicate 95% confidence intervals.

Figure A.5: Beliefs about the performance of different groups (Caste Experiment)

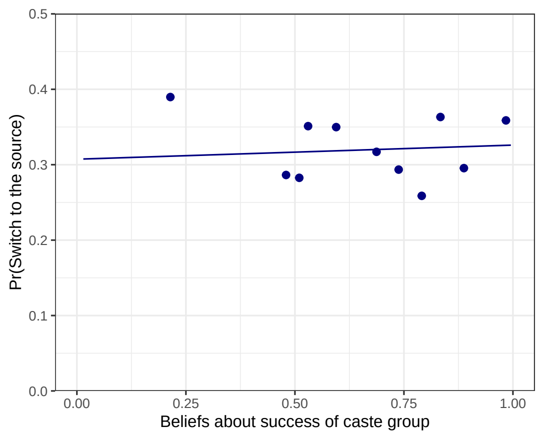


Notes. Cumulative frequencies of participant beliefs about the likelihood that a person from a particular caste category or religion makes an accurate first-period estimate in the task. Red line: O category Hindu, Black line: G category Hindu.

Figure A.6: Role of underlying beliefs



(a) Quality of information is unknown



(b) Quality of information is known

Notes. Binned scatter plots of participants' likelihood of switching to a source (with an experimentally assigned caste identity) against incentivised beliefs about the likelihood that a random individual of that caste is likely to make a correct decision on the task. (a): *Caste - No Signal*, participants are not informed about the quality of the source. (b): *Caste*, participants are given the quality signal.

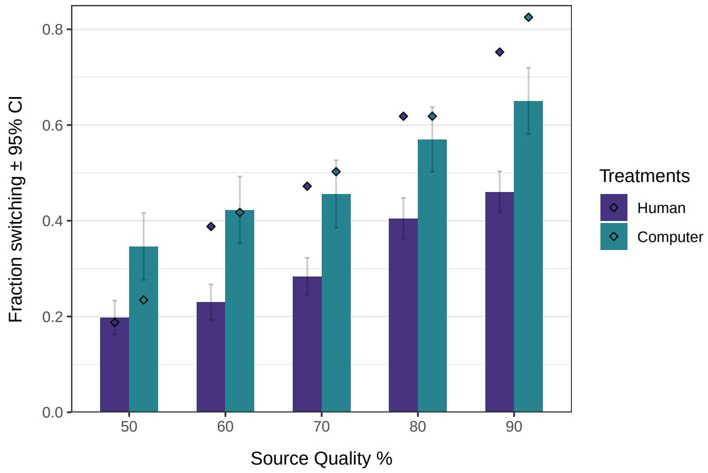
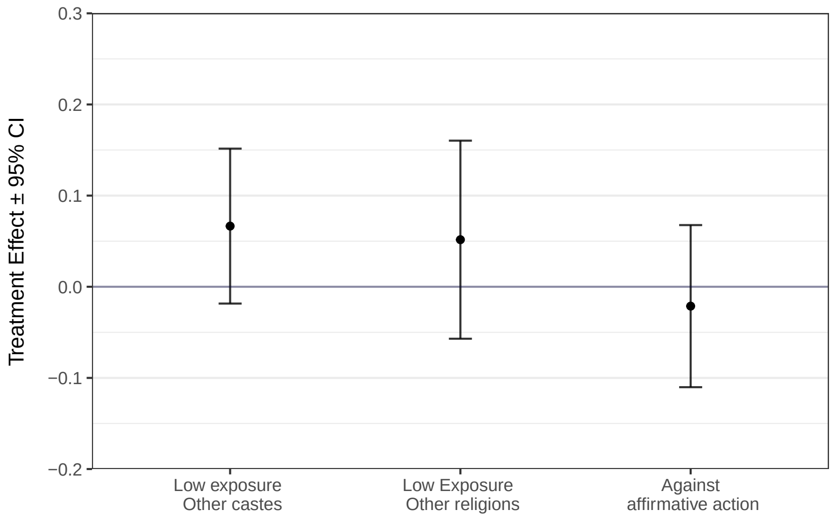


Figure A.7: Fraction of decisions where participants choose to switch to the source in treatments *Human* and *Computer*, by the quality of information. The diamonds indicate the fraction of decisions where switching would have been the optimal choice (based on a comparison of information quality and subjective certainty). Error bars indicate 95% confidence intervals.

Figure A.8: Mechanisms: *Caste* Experiment



Notes. Results from regressions of whether participants switch on an indicator of whether participants see an in-group source, an indicator of caste identity strength, and their interaction, on switching. The figure shows the estimated coefficients of the interaction term of whether participants see an in-group source and an indicator for (i) low exposure to people from other castes, (ii) low exposure to people from other religions, and (iii) whether they oppose caste-based affirmative action policies in the *Caste* experiment. Standard errors are clustered at the participant level in all regressions.

B Additional Tables

Table B.1: List of balls and urns tasks

#	Base Rate (p)	Red Balls (θ)	Total (Red) Draws	True Value (y_T)	Correct	Incorrect
1	0.7	70	3 (1)	50	51	25
2	0.7	70	5 (2)	50	48	26
3	0.9	70	3 (2)	95.5	95	76
4	0.9	90	3 (0)	1.2	3	23
5	0.9	70	5 (4)	99.1	97	82
6	0.5	70	5 (4)	92.7	92	77

Notes. Base rate is the probability with which the red bag is selected. *Red balls* is the number of red balls in the red bag. *Total (red) draws* is the number of balls that are drawn from the selected bag, with the number of red balls drawn in brackets. *True value* is the Bayesian posterior probability that the red bag was selected. *Correct* and *Incorrect* values for the source in each task are selected from responses on the same tasks in a previous study using the procedure described in Section 2.

Table B.2: Sample Descriptives

	Caste	Religion	Minimal	Human v Computer
Age	32	31	32	32
% Men	52	60	60	55
% College	90	89	77	75
% Employed	84	83	68	72
Decision Time	25	20	17	18
Participants	851	853	278	315
Decisions	5106	5118	1668	1890

Notes. Sample descriptives for the experiments presented in the paper. College and Employment are indicator variables = 1 if the participant has tertiary education or is not unemployed. Decision time is the average time (in seconds) taken by participants in the second period.

Table B.3: Regression analysis: Average treatment effects – Caste

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
General Caste Source	-0.028	-0.029	-0.028	-0.028	-0.027	-0.033
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Source is wrong			-0.041***	-0.041***	-0.015	-0.012
			(0.015)	(0.015)	(0.017)	(0.016)
General Caste				-0.015	-0.014	-0.015
				(0.022)	(0.022)	(0.022)
Quality					0.134***	0.140***
					(0.039)	(0.038)
Confidence						-0.003***
						(0.000)
Constant	0.334***	0.313***	0.325***	0.328***	0.227***	0.431***
	(0.015)	(0.023)	(0.060)	(0.060)	(0.068)	(0.072)
Task controls		✓	✓	✓	✓	✓
Demog. controls			✓	✓	✓	✓
R^2	0.001	0.003	0.008	0.009	0.011	0.035
Dependent variable mean	0.320	0.320	0.320	0.320	0.320	0.320
Observations	4,952	4,952	4,952	4,952	4,952	4,952
Individuals	851	851	851	851	851	851

Notes. Average treatment effects in experiment *Caste*, estimates from OLS regressions of Equation 2. Controls are task and order fixed effects, whether the source is correct, and demographic characteristics (the participant's age and dummies for employment, college education and gender). Robust standard errors, clustered at the participant level. $*\,p < 0.10, **\,p < 0.05, ***\,p < 0.01$.

Table B.4: Regression analysis: Average treatment effects – Religion

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Hindu Source	-0.010	-0.010	-0.008	-0.007	-0.007	-0.006
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Source is wrong			-0.050***	-0.050***	-0.039***	-0.038***
			(0.013)	(0.013)	(0.014)	(0.014)
Hindu				-0.045*	-0.046*	-0.042*
				(0.023)	(0.023)	(0.023)
Quality					0.130***	0.127***
					(0.046)	(0.045)
Confidence						-0.002***
						(0.000)
Constant	0.241***	0.228***	0.314***	0.341***	0.246***	0.410***
	(0.014)	(0.020)	(0.054)	(0.056)	(0.066)	(0.072)
Task controls		✓	✓	✓	✓	✓
Demog. controls			✓	✓	✓	✓
R^2	0.000	0.003	0.008	0.010	0.012	0.027
Dependent variable mean	0.236	0.236	0.236	0.236	0.236	0.236
Observations	5,063	5,063	5,063	5,063	5,063	5,063
Individuals	853	853	853	853	853	853

Notes. Average treatment effects in experiment *Religion*, estimates from OLS regressions of Equation 2. Controls are task and order fixed effects, whether the source is correct, and demographic characteristics (the participant's age and dummies for employment, college education and gender). Robust standard errors, clustered at the participant level. $* \, p < 0.10, ** \, p < 0.05, *** \, p < 0.01$.

Table B.5: Regression analysis: Average treatment effects – Minimal identity

	Dep. var.: Switch to source				
	(i)	(ii)	(iii)	(iv)	(v)
In-group Source	-0.018	-0.018	-0.025	-0.030	-0.010
	(0.028)	(0.029)	(0.028)	(0.029)	(0.027)
Source is wrong			-0.110***	-0.049*	-0.049*
			(0.025)	(0.026)	(0.026)
Quality				0.606***	0.614***
				(0.092)	(0.092)
Confidence					-0.005***
					(0.001)
Constant	0.321***	0.361***	0.487***	0.047	0.241**
	(0.019)	(0.039)	(0.079)	(0.105)	(0.101)
Task controls		✓	✓	✓	✓
Demog. controls			✓	✓	✓
R^2	0.000	0.005	0.023	0.052	0.113
Dependent variable mean	0.312	0.312	0.312	0.312	0.312
Observations	1,629	1,629	1,629	1,629	1,629
Individuals	278	278	278	278	278

Notes. Average treatment effects in experiment *Minimal*, estimates from OLS regressions of Equation 2. Controls are task and order fixed effects, whether the source is correct, and demographic characteristics (the participant's age and dummies for employment, college education and gender). Robust standard errors, clustered at the participant level. $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$.

Table B.6: Regression analysis: Average treatment effects – Human vs. Computer

	Dep. var.: Switch to source Pooling *Minimal*					Excluding *Minimal*
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Computer Source	0.173***	0.173***	0.174***	0.176***	0.156***	0.151***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	(0.025)
Source is wrong			-0.121***	-0.058***	-0.061***	-0.074***
			(0.017)	(0.018)	(0.018)	(0.024)
Quality				0.651***	0.642***	0.668***
				(0.064)	(0.063)	(0.088)
Confidence					-0.005***	-0.005***
					(0.000)	(0.000)
Constant	0.317***	0.376***	0.479***	0.002	0.212***	0.203**
	(0.012)	(0.026)	(0.056)	(0.073)	(0.071)	(0.101)
Task controls		✓	✓	✓	✓	✓
Demog. controls			✓	✓	✓	✓
R^2	0.026	0.032	0.047	0.079	0.134	0.142
Dependent variable mean	0.365	0.365	0.365	0.365	0.365	0.412
Observations	3,473	3,473	3,473	3,473	3,473	1,844
Individuals	593	593	593	593	593	315

Notes. Average treatment effects in experiment *Human vs. Computer*, estimates from OLS regressions of Equation 2. Controls are task and order fixed effects, whether the source is correct, and demographic characteristics (the participant's age and dummies for employment, college education and gender). Robust standard errors, clustered at the participant level. $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$.

Table B.7: Regression analysis: Experiment *Caste – No Signal*.

	Dep. var.: Switch to source			
	(i)	(ii)	(iii)	(iv)
General Caste Source	-0.025	-0.021	-0.021	-0.021
	(0.034)	(0.034)	(0.034)	(0.034)
Quality		0.034	0.019	0.019
		(0.022)	(0.024)	(0.024)
Confidence		-0.002***	-0.002***	-0.002***
		(0.001)	(0.001)	(0.001)
Source is wrong			-0.035	-0.035
			(0.025)	(0.025)
General Caste				0.000
				(0.038)
Constant	0.285***	0.452***	0.474***	0.474***
	(0.025)	(0.105)	(0.104)	(0.104)
Task controls		✓	✓	✓
Demog. controls		✓	✓	✓
R^2	0.001	0.031	0.032	0.032
Dependent variable mean	0.273	0.273	0.273	0.273
Observations	1,719	1,719	1,719	1,719
Individuals	295	295	295	295

Notes. Average treatment effects in the Caste – No Signal experiment. In this experiment, the quality signal is not provided to participants. Controls are task and order fixed effects, and demographic characteristics(the participant's age and dummies for employment, college education and gender). Standard errors are clustered at the participant level. $* p < 0.10, ** p < 0.05, ***p < 0.01$.

Table B.8: Robustness – In-group preferences, *Caste* experiment

	Dep. var.: Switch to source					
	Full sample	First 2 tasks	No Speeders	No speeders or slackers	Atleast 1 switch	Same Task <= 10
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
General Caste Source	-0.013	0.021	-0.011	-0.012	-0.024	-0.020
	(0.035)	(0.042)	(0.035)	(0.039)	(0.036)	(0.051)
General Caste	0.000	0.007	-0.009	-0.028	-0.009	0.015
	(0.031)	(0.036)	(0.032)	(0.035)	(0.033)	(0.049)
General Caste Source × General Caste	-0.032	-0.069	-0.025	-0.016	-0.037	0.004
	(0.043)	(0.053)	(0.044)	(0.048)	(0.045)	(0.067)
Quality	0.056***	0.063**	0.068***	0.062***	0.152***	0.127**
	(0.015)	(0.026)	(0.017)	(0.020)	(0.050)	(0.059)
Constant	0.490***	0.500***	0.484***	0.546***	0.559***	0.508***
	(0.068)	(0.084)	(0.072)	(0.082)	(0.081)	(0.116)
Task controls	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓
R^2	0.035	0.051	0.039	0.042	0.029	0.043
Dependent variable mean	0.320	0.324	0.325	0.311	0.460	0.291
Observations	4,952	1,653	4,008	3,053	3,444	2,055
Individuals	851	851	842	825	590	354

Notes. OLS estimates of Equation 3 for the robustness analysis of in-group preferences in the Caste experiment. Controls are task and order fixed effects, and demographic characteristics(the participant's age and dummies for employment, college education and gender). Standard errors are clustered at the participant level. $* p < 0.10, ** p < 0.05, ***p < 0.01$.

Table B.9: Robustness – In-group preferences, *Religion* experiment

	Dep. var.: Switch to source				
	Full sample	First 2 tasks	No Speeders	No speeders or slackers	Atleast 1 switch
	(i)	(ii)	(iii)	(iv)	(v)
Hindu Source	0.079	0.027	0.077	0.071	0.157**
	(0.065)	(0.081)	(0.068)	(0.073)	(0.075)
Hindu Source × Hindu	-0.093	-0.013	-0.090	-0.079	-0.159**
	(0.068)	(0.085)	(0.071)	(0.077)	(0.080)
Hindu	0.003	-0.025	-0.001	0.000	0.053
	(0.041)	(0.054)	(0.043)	(0.047)	(0.049)
Quality	0.091*	0.103	0.139***	0.136**	0.174**
	(0.049)	(0.079)	(0.052)	(0.059)	(0.077)
Constant	0.372***	0.409***	0.370***	0.403***	0.332***
	(0.085)	(0.116)	(0.088)	(0.096)	(0.111)
Task controls	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓
R^2	0.024	0.020	0.029	0.035	0.025
Dependent variable mean	0.230	0.231	0.236	0.225	0.407
Observations	4,348	1,449	3,698	2,873	2,458
Individuals	733	733	732	724	414

Notes. OLS estimates of Equation 3 for the robustness analysis of in-group preferences in the Religion experiment. Controls are task and order fixed effects, and demographic characteristics(the participant's age and dummies for employment, college education and gender). Standard errors are clustered at the participant level. $* p < 0.10, ** p < 0.05, *** p < 0.01$.

Table B.10: Regression analysis: Mechanisms in the Caste experiment

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
General Caste Source	-0.028	-0.036	0.014	-0.024	-0.025	-0.031
	(0.053)	(0.022)	(0.069)	(0.021)	(0.025)	(0.022)
General Caste Source × Quality	-0.007					
	(0.071)					
General Caste Source × Source is wrong		0.010				
		(0.030)				
General Caste Source × Confidence			0.001			
			(0.001)			
General Caste Source × Guess 1 correct				-0.061*		
				(0.037)		
General Caste Source × Dist. to Source					0.000	
					(0.001)	
General Caste Source × Decision time (z)						0.008
						(0.017)
Quality	0.144***	0.141***	0.145***	0.142***	0.142***	0.148***
	(0.054)	(0.038)	(0.039)	(0.038)	(0.038)	(0.040)
Source is wrong	-0.012	-0.016	-0.010	-0.008	-0.011	-0.007
	(0.016)	(0.022)	(0.016)	(0.016)	(0.016)	(0.017)
Confidence	-0.003***	-0.003***	-0.004***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Guess 1 correct				-0.071**		
				(0.029)		
Dist. to Source					0.000	
					(0.000)	
Decision time (z)						0.021
						(0.014)
Constant	0.424***	0.428***	0.426***	0.414***	0.406***	0.447***
	(0.076)	(0.071)	(0.081)	(0.071)	(0.073)	(0.077)
Task controls	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓
R^2	0.035	0.035	0.039	0.040	0.035	0.038
Dependent variable mean	0.320	0.320	0.320	0.320	0.320	0.312
Observations	4,952	4,952	4,907	4,952	4,952	4,346
Individuals	851	851	850	851	851	843

Notes. The table presents estimates from OLS regressions of switching on the treatment indicator interacted with one of six mechanism variables within the *Caste* experiment, controlling for demographics, task parameters, and task order. Demographic controls are the participant's age and dummies for employment, college education and gender. Standard errors are clustered at the participant level. *Quality* is a continuous variable ranging from 0.5 to 0.9. *Source is wrong*=1 if the shown source's estimate is incorrect. Confidence is a participant's confidence in the accuracy of their independent estimate (0 to 100). *Guess 1 correct*= 1 if the participant made a correct first period estimate. *Dist. to source* $= |y_s - y_1|$. *Decision time* is the z-standardised time in seconds taken by participant on the page where they decide whether to stick or switch. $* \ p < 0.10, ** \ p < 0.05, *** p < 0.01$.

Table B.11: Regression analysis: Mechanisms in the Religion experiment

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Hindu Source	0.164**	-0.004	-0.009	0.002	-0.009	0.001
	(0.068)	(0.023)	(0.068)	(0.021)	(0.022)	(0.021)
Hindu Source × Quality	-0.238**					
	(0.092)					
Hindu Source × Source is wrong		0.004				
		(0.028)				
Hindu Source × Confidence			0.000			
			(0.001)			
Hindu Source × Guess 1 correct				-0.035		
				(0.036)		
Hindu Source × Dist. to Source					0.000	
					(0.001)	
Hindu Source × Decision time (z)						-0.003
						(0.017)
Quality	0.212***	0.088*	0.092*	0.090*	0.087*	0.072
	(0.068)	(0.049)	(0.049)	(0.049)	(0.049)	(0.052)
Source is wrong	-0.047***	-0.048**	-0.049***	-0.045***	-0.046***	-0.036**
	(0.015)	(0.021)	(0.015)	(0.015)	(0.015)	(0.016)
Confidence	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Guess 1 correct				-0.076***		
				(0.028)		
Dist. to Source					0.002***	
					(0.000)	
Decision time (z)						0.037***
						(0.014)
Constant	0.294***	0.382***	0.383***	0.377***	0.325***	0.367***
	(0.083)	(0.077)	(0.085)	(0.077)	(0.077)	(0.080)
Task controls	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓
R^2	0.023	0.021	0.021	0.026	0.033	0.029
Dependent variable mean	0.230	0.230	0.229	0.230	0.230	0.224
Observations	4,348	4,348	4,317	4,348	4,348	3,801
Individuals	733	733	733	733	733	729

Notes. The table presents estimates from OLS regressions of switching on the treatment indicator interacted with one of six mechanism variables within the *Religion* experiment, controlling for demographics, task parameters, and task order. Demographic controls are the participant's age and dummies for employment, college education and gender. Standard errors are clustered at the participant level. *Quality* is a continuous variable ranging from 0.5 to 0.9. *Source is wrong*=1 if the shown source's estimate is incorrect. Confidence is a participant's confidence in the accuracy of their independent estimate (0 to 100). *Guess 1 correct*= 1 if the participant made a correct first period estimate. *Dist. to source* $= |y_s - y_1|$. *Decision time* is the z-standardised time in seconds taken by participant on the page where they decide whether to stick or switch. $* \, p < 0.10, ** \, p < 0.05, *** p < 0.01$.

Table B.12: Regression analysis: Mechanisms in the Minimal experiment

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
In-group Source	0.070	-0.021	-0.039	-0.005	-0.009	-0.011
	(0.120)	(0.032)	(0.078)	(0.029)	(0.034)	(0.027)
In-group Source × Quality	-0.113					
	(0.172)					
In-group Source × Source is wrong		0.037				
		(0.048)				
In-group Source × Confidence			0.000			
			(0.001)			
In-group Source × Guess 1 correct				-0.008		
				(0.053)		
In-group Source × Dist. to Source					0.000	
					(0.001)	
In-group Source × Decision time (z)						0.015
						(0.027)
Quality	0.673^{***}	0.611^{***}	0.621^{***}	0.616^{***}	0.615^{***}	0.557^{***}
	(0.130)	(0.092)	(0.092)	(0.090)	(0.091)	(0.095)
Source is wrong	-0.048^{*}	-0.068^{*}	-0.050^{*}	-0.047^{*}	-0.052^{**}	-0.055^{**}
	(0.026)	(0.035)	(0.026)	(0.026)	(0.026)	(0.026)
Confidence	-0.005^{***}	-0.005^{***}	-0.004^{***}	-0.004^{***}	-0.005^{***}	-0.005^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Guess 1 correct				-0.131^{***}		
				(0.038)		
Dist. to Source					0.002^{**}	
					(0.001)	
Decision time (z)						0.057^{***}
						(0.021)
Constant	0.200	0.249^{**}	0.250^{**}	0.232^{**}	0.190^{*}	0.210^{**}
	(0.126)	(0.100)	(0.107)	(0.100)	(0.102)	(0.106)
Task controls	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓
R^2	0.113	0.113	0.109	0.122	0.120	0.136
Dependent variable mean	0.312	0.312	0.308	0.312	0.312	0.300
Observations	1,629	1,629	1,613	1,629	1,629	1,461
Individuals	278	278	278	278	278	277

Notes. The table presents estimates from OLS regressions of switching on the treatment indicator interacted with one of six mechanism variables within the *Minimal* experiment, controlling for demographics, task parameters, and task order. Demographic controls are the participant's age and dummies for employment, college education and gender. Standard errors are clustered at the participant level. *Quality* is a continuous variable ranging from 0.5 to 0.9. *Source is wrong*=1 if the shown source's estimate is incorrect. Confidence is a participant's confidence in the accuracy of their independent estimate (0 to 100). *Guess 1 correct*= 1 if the participant made a correct first period estimate. *Dist. to source* $= |y_s - y_1|$. *Decision time* is the z-standardised time in seconds taken by participant on the page where they decide whether to stick or switch. $* \, p < 0.10, ** \, p < 0.05, *** p < 0.01$.

Table B.13: Regression analysis: Mechanisms in the Human v Computer experiment

	Dep. var.: Switch to source					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Computer Source	0.126	0.149***	0.176***	0.159***	0.196***	0.156***
	(0.094)	(0.025)	(0.051)	(0.022)	(0.028)	(0.022)
Computer Source × Quality	0.043					
	(0.135)					
Computer Source × Source is wrong		0.022				
		(0.040)				
Computer Source × Confidence			0.000			
			(0.001)			
Computer Source × Guess 1 correct				-0.008		
				(0.051)		
Computer Source × Dist. to Source					-0.002**	
					(0.001)	
Computer Source × Decision time (z)						-0.044**
						(0.019)
Quality	0.630***	0.641***	0.642***	0.643***	0.644***	0.647***
	(0.072)	(0.063)	(0.063)	(0.062)	(0.063)	(0.066)
Source is wrong	-0.062***	-0.067***	-0.061***	-0.059***	-0.063***	-0.059***
	(0.018)	(0.019)	(0.018)	(0.017)	(0.018)	(0.018)
Confidence	-0.005***	-0.005***	-0.005***	-0.004***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Guess 1 correct				-0.146***		
				(0.022)		
Dist. to Source					0.003***	
					(0.001)	
Decision time (z)						0.051***
						(0.012)
Constant	0.221***	0.214***	0.207***	0.205***	0.157**	0.155**
	(0.075)	(0.071)	(0.071)	(0.071)	(0.072)	(0.076)
Task controls	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓
R^2	0.134	0.135	0.135	0.145	0.141	0.146
Dependent variable mean	0.365	0.365	0.365	0.365	0.365	0.358
Observations	3,473	3,473	3,473	3,473	3,473	3,130
Individuals	593	593	593	593	593	592

Notes. The table presents estimates from OLS regressions of switching on the treatment indicator interacted with one of six mechanism variables within the *Computer* experiment, controlling for demographics, task parameters, and task order. Demographic controls are the participant's age and dummies for employment, college education and gender. Standard errors are clustered at the participant level. *Quality* is a continuous variable ranging from 0.5 to 0.9. *Source is wrong*=1 if the shown source's estimate is incorrect. Confidence is a participant's confidence in the accuracy of their independent estimate (0 to 100). *Guess 1 correct*= 1 if the participant made a correct first period estimate. *Dist. to source* $= |y_s - y_1|$. *Decision time* is the z-standardised time in seconds taken by participant on the page where they decide whether to stick or switch. $*\, p < 0.10, **\, p < 0.05, ***p < 0.01$.

C Experiment details and additional analyses

C.1 Survey module

Demographics. In all experiments, participants provide their age, sex, education level, and employment status. In the *Caste* and *Religion* experiments, participants also state their religion, caste group, religiosity, and their favourite religious festival (free text). These questions also serve as a mild priming device, drawing attention to these characteristics before the main decision tasks (as in Chen et al. (2014)). Participants are only asked these questions after clearing the attention and comprehension tests.

Exposure and Attitudes. Participants in the *Caste* experiments respond to three additional survey questions after they complete the main tasks. The first two elicit the extent to which respondents have close associations with people who belong to their religious or caste out-groups. The questions are:

> How many of your friends belong to the same [Religion/Caste category] as you?

The final question elicits people's attitudes towards caste-based affirmative action policies.

> Do you support reservations in jobs and educational institutions based on caste?

Beliefs – Group performance. Participants in the *Caste* experiments state their beliefs about the probability that an anonymous person belonging to a caste or religious group will answer the experimental task correctly. These beliefs are incentivised for accuracy – participants earn an additional $0.50 if they guess the number within ±5% points of the true probability, which is calculated based on a previous study. These questions are asked for the "General" and "Scheduled Castes" caste groups.

> If a randomly selected individual belonging to the [Caste group] category attempted the Decision task (the task that you just completed). What do you believe is the probability (0% to 100%) that they will answer it correctly? 0% means that they will never get it correct. 100% means that they will always get it correct.

Reflection question. After completing the survey, respondents provide free text responses to:

> Please tell us how you used the recommendation when making the second decision in the tasks. How did you think about the choice of using your own decision or the shown number?

Individualism. Experiments *Religion, Minimal*, and *Human v. Computer* included 8 additional questions that measured Horizontal and Vertical Individualism. These measures were taken from the scale developed by Triandis and Gelfand (1998).[16] Answers are coded on a scale from 1 to 5, with 1 indicating "Strongly disagree" and 5 indicating "Strongly agree".

The questions are:

I would rather depend on myself than others.
I rely on myself most of the time. I rarely rely on others.
I often do my own thing.
My personal identity, independent of others, is very important to me.
It is important that I do my job better than others.
Winning is everything.
Competition is the law of nature.
When another person does better than I do, I get tense.

Summing up the responses to these questions results in an "Individualism" score for each participant. I preregistered analyses studying heterogeneity in switching by this individualism measure. The results do not indicate any significant correlations between individualism and the likelihood of switching because of differences in the source's identity.

C.2 Surnames used in the experiments

The recognisability of the Hindu surnames used in the experiments were validated in a separate survey ($N = 350$). These individuals completed much of the training module and comprehension tests that will be used in the main treatments. The source of truth for the incentivisation of these classifications came from official classifications of individuals belonging to these communities, or the common nature of these surnames. The recognisability ranged from $\approx$ 60% to 85%, with names from the SC/ST/OBC castes being, on average, more recognisable than General caste surnames. The Muslim names were not validated, given the common nature and clear identifiability of these names.

Hindu, General caste names. Iyer, Banerjee, Chaturvedi, Tiwari, Bharadwaj, Mishra.

Hindu, SC/ST/OBC caste names. Paraiyar, Bhil, Jatav, Manjhi, Mahar, Chamar.

Muslim names. Khan, Shaikh, Abdullah, Syed, Moinuddin, Ali.

At the start of the experiment, participants are told that the source's decisions were made by participants in a previous study, which was a true statement. Participants were informed at the end of the study that the names used in Religion and Caste experiments were nicknames, and not the actual names of the people who made the decisions.

C.3 Source decisions

I curate the decisions as follows:

[16]The full battery comprises 24 questions, half of which measure individualism and the other half collectivism. In the experiment, the questions regarding collectivism are dropped to avoid making the survey module too lengthy, focusing on the role between individualism and a preference for consistency or autonomy.

- For each task, two decisions are selected such that at least one participant from each relevant identity group made that decisions.
- For each task, one of the decisions is within ±2% of the Bayesian posterior. This is the "correct" answer.
- The second decisions ("incorrect") is chosen to be at least 15% points away from the true value.

For example, 51 and 25 are chosen on task 1. At least one participant from each identity category made these decisions in the pilot experiments.

C.4 Quality signals

Conceptual discussion

The DM forms a belief $\hat{Q}$ about the accuracy Q of y_s. This is the probability that $y_s = y_T$. The DM also has a belief about the accuracy of their own decision, their confidence c, which is the probability that $y_1 = y_T$. Suppose that choosing y_T yields a utility $U = 1$, and that all other choices yield $U = 0$. This gives:

$$E[\text{switch}] = \hat{Q} \cdot 1 - (1 - \hat{Q}) \cdot 0 = \hat{Q}$$
$$E[\text{stick}] = c \cdot 1 - (1 - c) \cdot 0 = c$$

When deciding whether to stick or switch, the only available comparison is between c and $\hat{Q}$, i.e. the DM switches if $E[\text{switch}] > E[\text{stick}]$. This implies that DMs switch if they believe that $\hat{Q} > c$, and stick otherwise.

$$y_2 = \begin{cases} y_s & \text{if } \hat{Q} > c \\ y_1 & \text{if } \hat{Q} \leq c \end{cases}$$

First, consider the case where a DM only sees y_s and has no information about the source's accuracy. If the DM has no prior beliefs about the pool from which sources come from, $\hat{Q} = 1/n$ which means they will only switch if they are very uncertain about their decision.

Now suppose that the social identity (such as a person's race, ethnicity, gender, or political affiliation) g that S belongs to is observable. Models of statistical discrimination show that people make inferences about individuals from beliefs about the characteristics of their social groups. In this case, if the DM has some prior beliefs about the ability of an identity group g they can evaluate the accuracy of the decision y_s based on these beliefs, i.e. $\hat{Q} = Q_g$. Thus, DMs reactions to sources of different identities will be driven by differences in the beliefs held about these groups.[17]

Suppose now that the DM receives a (possibly noisy) signal about the accuracy Q_S of the specific decision y_s made by the source S. The DM's belief about the accuracy of y_s will be a function of their

[17] If these beliefs are inaccurate, then this will lead to behaviour resembling inaccurate statistical discrimination (Bohren et al., 2023).

beliefs about the group g and the signal, $\hat{Q} = f(Q_g, Q_S)$. A Bayesian updating rule would predict that these are weighted in proportion to the (beliefs in the) variance of Q_g and Q_S. If Q_S is sufficiently precise then $\hat{Q} = Q_S = Q$, and underlying beliefs about the ability of groups will not play a role in decision-making. In other words, providing a precise signal of the quality of the decision made by the source will eliminate the role of any group-specific beliefs in this situation. This yields the following predictions when a signal of quality is provided:

Prediction. *Switching increases with an increase in beliefs about information quality (Q_S). When Q_S is precise, switching is uncorrelated with Q_g.*

Given a precise signal, any observed differences in switching because of a difference in source's identity can be attributed to differences in k because of different group identities, which can be interpreted as preferences for one group relative to another. The signal essentially converts the switching decision into a lottery, and breaks the link between beliefs about information quality and preferences for the identity of the source.

Instructions

In experiment *Caste*, the following information is shown to participants:

> There is a [Q] % chance that this number is the correct answer.
>
> The computer has access to a pool of participants who made correct guesses on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess. The computer randomly chooses one of these people and shows you a number:
>
> - With [Q]% probability, the shown number is the chosen person's correct guess.
> - Otherwise, the shown number is incorrect.

In the *Computer* treatment, the following information is shown to participants:

> There is a [Q] % chance that this value is within ±2% points of the correct answer.
>
> Otherwise, the computer chooses a random number between 0 and 100.

In all other experiments, the following information is shown to participants:

> There is a [Q] % chance that this guess is within ±2% points of the correct answer.
>
> A computer randomly chooses this guess from the guesses made on this exact task by participants in a previous study - they had the same information, and saw the same ball colours when making their guess.

In all of the above cases, Q is a placeholder for the randomly chosen probability that the source's decision is correct (i.e. participants always see the probability).

C.5 Experiment *Caste – No signal*

This experiment was conducted at the same time and participants were recruited from the same sampling pool as the *Caste* experiment. Participants were given a quality signal in the *Caste* experiment, but not in *Caste – No Signal*.

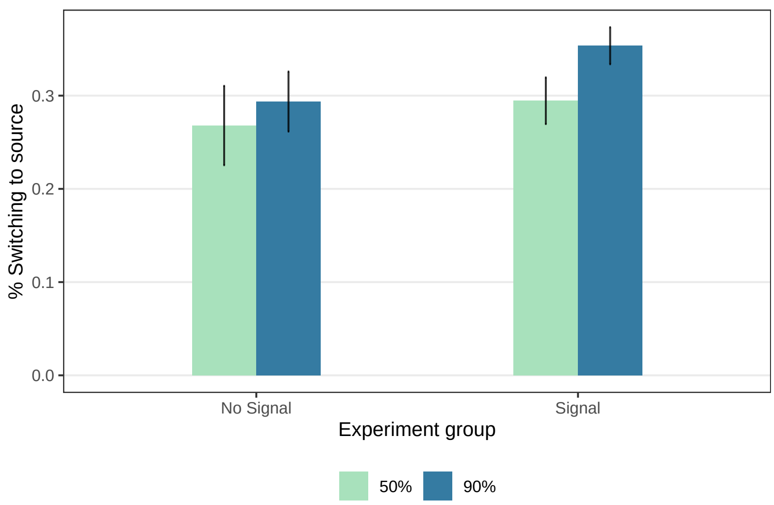
The following information is presented to participants in this experiment group instead of the quality signal:

> The computer has access to a pool of participants who made guesses on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess. The computer randomly chooses one of these people and shows you the chosen person's guess.

Learning with and without a signal

A comparison of decisions in these experiments shows that participants switch more when given the quality signal. The difference is more pronounced when the signal is stronger – Appendix Figure C.1 shows that when source quality is low (50%), switching in the two experiments differs by ≈ 11%. When the source quality is high (90%), the difference is ≈ 22%.

Figure C.1: Switching with and without the quality signal



Notes. The bars shows the percentage of participants switching to the source at different levels of signal quality. In the first group, signal quality is not revealed to participants (experiment *Caste – No Signal*). In the second group, participants are informed of the signal quality (experiment*Caste*). Error bars indicate 95% confidence intervals.

D Experiment instructions

Experiment *Caste – Quality*

Welcome

This study is being conducted by researchers from the Norwegian School of Economics. It will take about 10 minutes to complete this study. Please read all questions and instructions carefully.

Payment and Bonus Rewards

- You can earn up to **$ 3.5 (~ INR 290)** in additional bonuses depending on your task performance. You will be given information about the bonuses as you progress through the study.
- If you want to be eligible for the bonus, you must complete the entire study. The bonus that you earn will be shown to you at the end of the study, and will be processed and sent to you within 7 working days.
- You will need to read all the instructions carefully and answer a comprehension test correctly in order to participate in the study and be eligible for bonuses. You will have two chances to complete the comprehension test. In case you do not pass the comprehension test, you will not be eligible for any rewards.
- Bonus rewards will be paid in points, equivalent to the US dollar amount.

Guidelines

- We encourage you to try to answer all questions accurately and truthfully.
- This study is confidential and will only be used for research purposes.
- This study has been approved by an institutional review board.
- You may write to us at s14117@nhh.no in case you have any queries about this study.

Consent

By participating in this study you agree to the usage of your anonymised information and actions within the survey for research purposes.

Next

Figure D.1: Welcome and consent

Instructions for the Decision Tasks (1/3)

You will do 6 of these tasks, and in each task, you will make two decisions. The instructions below apply to all tasks.

Task Description

In each task, there are two bags - a **RED** bag and a **BLACK** bag. Each bag contains 100 balls, which are either **Red** or **Black**. The **RED** bag always contains more Red balls, and the **BLACK** bag contains more black balls. You know how many balls of each colour are in a bag.

One of the bags is chosen at random, but you do not know which bag is chosen. A few balls are drawn randomly from the chosen bag. Balls are drawn one by one, and they are put back into the bag before the next one is drawn.

You are shown the colours of all the balls that were drawn from the bag. You also know the chance with which a bag can be chosen, and the number of red and black balls in each bag.

You will make **two** decisions on each task. Both decisions are **eligible for a bonus reward.**

The First Decision

Based on the composition of the bags and the colours of the drawn balls, you have to **guess the likelihood that the balls were drawn from the RED bag. This should be a number between 0 and 100 %.** 0 means that you think there is no chance the balls were drawn from the red bag, and 100 means that you are completely certain that the balls were drawn from the red bag.

Next

Figure D.2: Instructions 1/3

Instructions for the Decision Tasks (2/3)

After making the first decision, you will be shown a number.

How is this number generated?

- The computer has access to a pool of participants who made **correct guesses** on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess.
- The computer randomly chooses one of these people, and shows you a number:
    - With some probability, the shown number is the chosen person's **correct** guess.
    - Otherwise, the shown number is **incorrect**.

The probability that the correct guess is shown can change from task to task.

The Second Decision

You have to make a choice:

1. Stick to your first guess, OR
2. Switch to the shown number.

Next, you will learn about how you can earn bonuses on these tasks.

Next

Figure D.3: Instructions 2/3

Instructions for the Decision Tasks (3/3)

Bonuses

- You will do 6 of these tasks, making 2 decisions in each task. **One** of these 12 decisions will be randomly selected for a bonus.
- Each task has a mathematically correct answer.
- If the chosen guess is **within ± 2 % points of the correct answer**, your bonus will be \$ 3.
- If your guess is more than 2 % away from the correct answer, then you will not earn a bonus.

It is therefore in your best interest to try and make all your guesses as accurately as possible.

Note: The correct answer to each task can be calculated using a probability formula. Click to see formula

Next, you will see an example task.

Next

Figure D.4: Instructions 3/3

Example Task (1/3)

This is an example of how the actual task will work.

Example Task

One of these bags is selected.
The **RED** bag is selected with a **50%** chance.

RED bag	**BLACK** bag
100 **Red**, 0 **Black** balls	0 **Red**, 100 **Black** balls
50% chance of being selected	50% chance of being selected

When you click the button, **3 balls** will be drawn from the selected bag one by one, with replacement. You will be allowed to make your guess after drawing the balls.

Click to draw balls

Figure D.5: Example 1/3

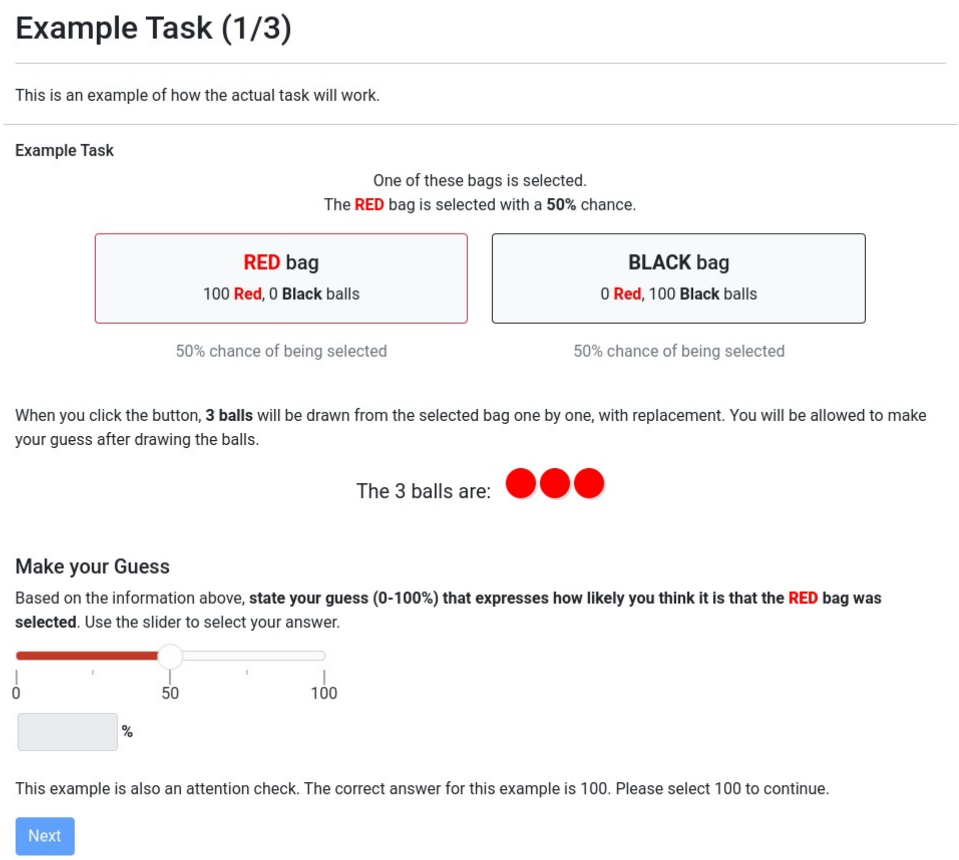


Figure D.6: Example 1/3

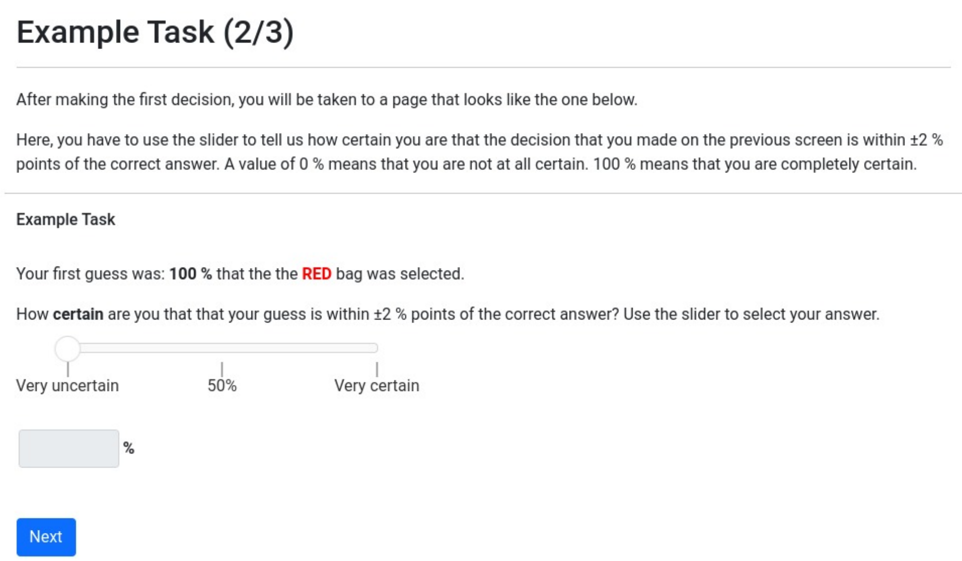


Figure D.7: Example 2/3

Example Task (3/3)

The example continues on this page - no bonuses will be awarded for this task, but you have to provide an answer to continue to the next page.

Example Task - The Second Decision

The computer shows:

Study Participant

73

There is a **50 % chance** that this number is the correct answer.

More info:

- The computer has access to a pool of participants who made **correct guesses** on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess.
- The computer randomly chooses one of these people, and shows you a number:
  - With some probability, the shown number is the chosen person's **correct** guess.
  - Otherwise, the shown number is **incorrect**.

Make your Second Decision

Your first guess was: **100 %**.

Based on the information above, **choose one of these options to make your decision (0-100%) that expresses how likely you think it is that the RED bag was selected**.

- **100 %** - Stick to your first guess
- **73 %** - Switch to the shown number

Next

Figure D.8: Example 3/3

Summary

Congratulations! You have successfully completed the example task.

Next, you will answer a short comprehension test. You must pass the test in order to be eligible for survey completion and bonus payments. Below is a short summary of the task, please familiarise yourself with them before moving on to the test.

Task Description

There are two bags - a **RED** bag and a **BLACK** bag. Each bag contains 100 balls, which are either **Red** or **Black**. The **RED** bag always contains more Red balls, and the **BLACK** bag contains more black balls. You know how many balls of each colour are in a bag.

One of the bags is chosen at random, and a few balls are drawn randomly from the chosen bag. Balls are drawn one by one, and they are put back into the bag before the next one is drawn. You are shown the colours of all the balls that were picked.

First decision: Based on this information, you have to **guess the likelihood that the balls were drawn from the RED bag. This should be a number between 0 and 100 %.**

Second decision: The second decision is made after you see what someone else has guessed on the exact same task. They have seen the same bags and drawn balls as you. A computer will show you a number. You have to decide whether to stick to your first guess, or switch to the shown number.

Important points

1. You make two decisions on each task, and there are 6 tasks.
2. Each of the Decision Tasks has a mathematically correct answer.
3. One of the decisions will be randomly chosen for a bonus. You will get a $ 3 bonus if the chosen decision is within 2% points of the correct answer.

Click the button to proceed to the comprehension test.

Next

Figure D.9: Summary

Task Comprehension: First Attempt

You must answer all questions to proceed. You have two chances to do this. If you are unable to pass, the assignment will end immediately and you will not be eligible for approval or for any bonus rewards.

Click here to review the instructions.

1. You will make several guesses in the tasks. What exactly are you guessing?

- ○ A number between 0-100, representing your guess about the chance that the RED bag was selected.
- ○ The probability of a green ball being drawn.

2. Which decision will be used to calculate the bonus?

- ○ The first decision
- ○ The last decision
- ○ One randomly chosen decision

3. If a decision is selected for the bonus, you will get the bonus ...

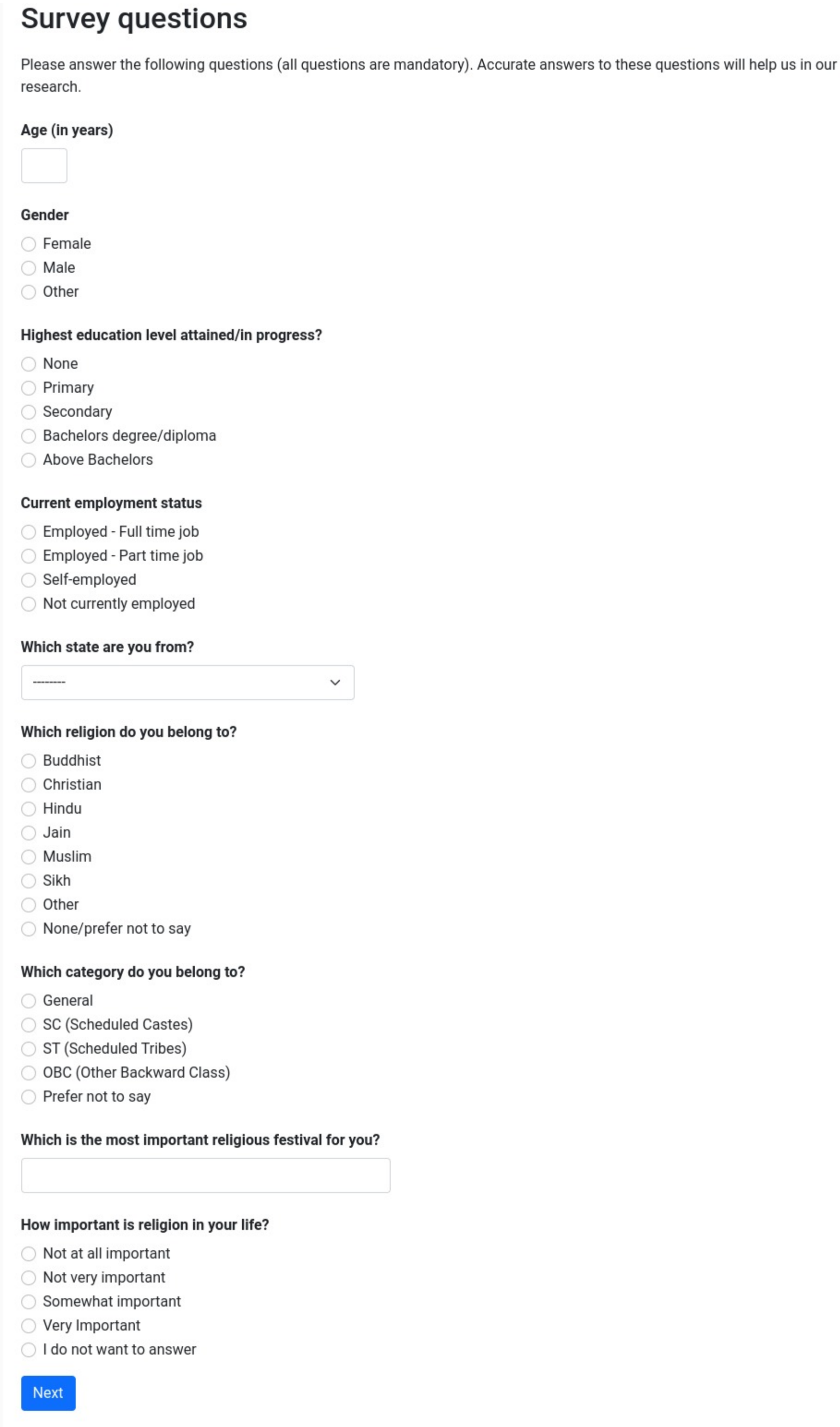
- ○ If the selected decision is within 2 % points of the mathematically correct value.
- ○ By correctly guessing the hard disk capacity of the computer.

4. Before making your second decision, the computer will show you a number. Which of these statements is correct?

- ○ The shown number is always a random number.
- ○ The shown number is either another person's correct guess, else it is an incorrect answer.

Next

Figure D.10: Comprehension (two attempts)



Survey questions

Please answer the following questions (all questions are mandatory). Accurate answers to these questions will help us in our research.

Age (in years)

Gender

- Female
- Male
- Other

Highest education level attained/in progress?

- None
- Primary
- Secondary
- Bachelors degree/diploma
- Above Bachelors

Current employment status

- Employed - Full time job
- Employed - Part time job
- Self-employed
- Not currently employed

Which state are you from?

Which religion do you belong to?

- Buddhist
- Christian
- Hindu
- Jain
- Muslim
- Sikh
- Other
- None/prefer not to say

Which category do you belong to?

- General
- SC (Scheduled Castes)
- ST (Scheduled Tribes)
- OBC (Other Backward Class)
- Prefer not to say

Which is the most important religious festival for you?

How important is religion in your life?

- Not at all important
- Not very important
- Somewhat important
- Very Important
- I do not want to answer

Next

Figure D.11: Demographics

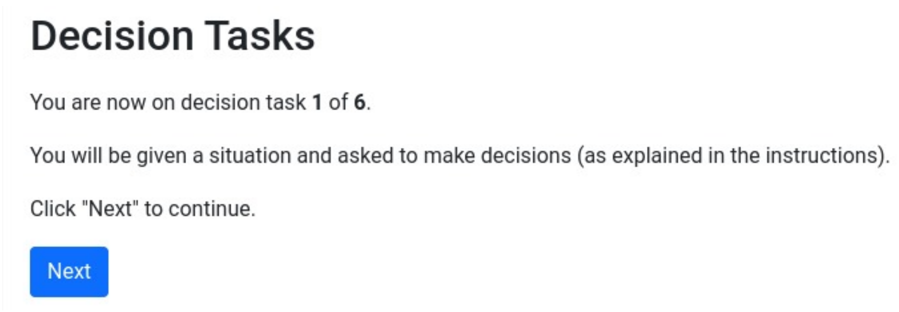


Figure D.12: In-between decision tasks

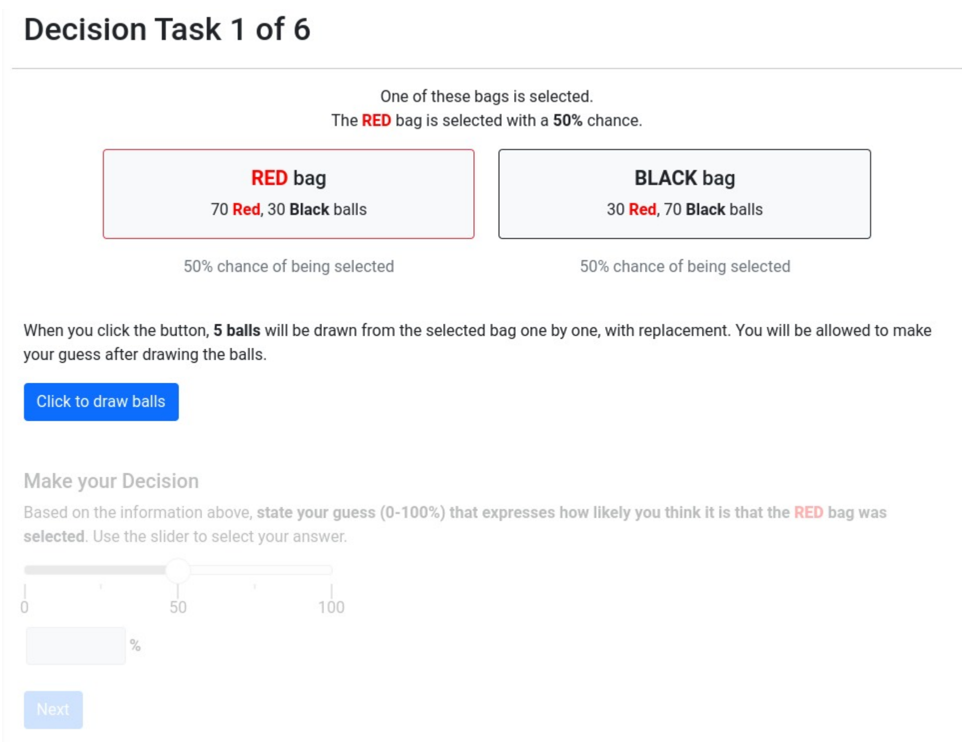


Figure D.13: Decision Tasks - First guess

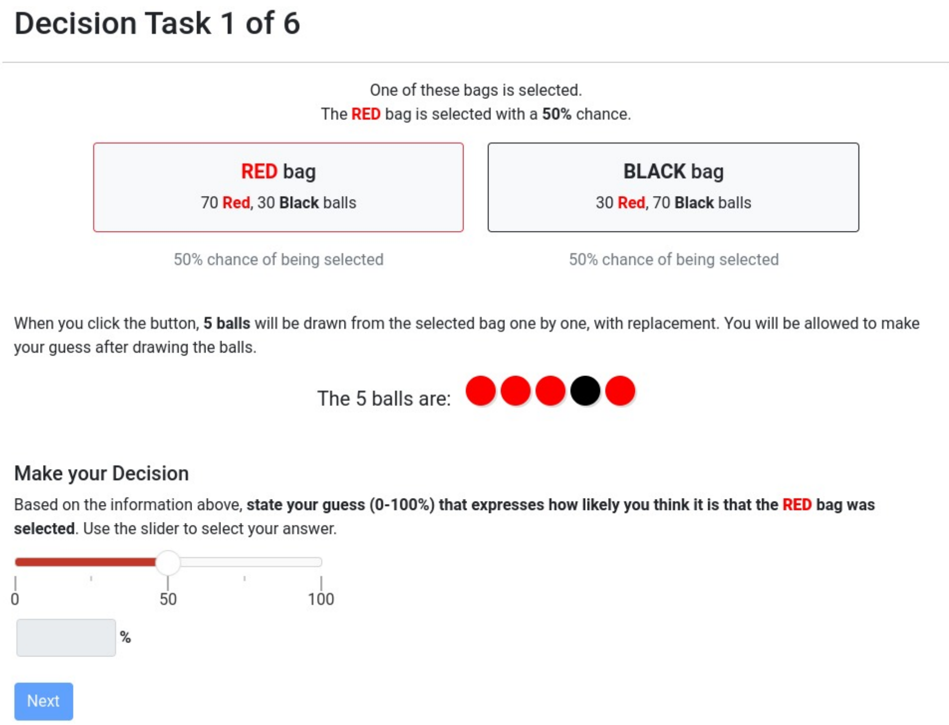


Figure D.14: Decision Tasks - First guess

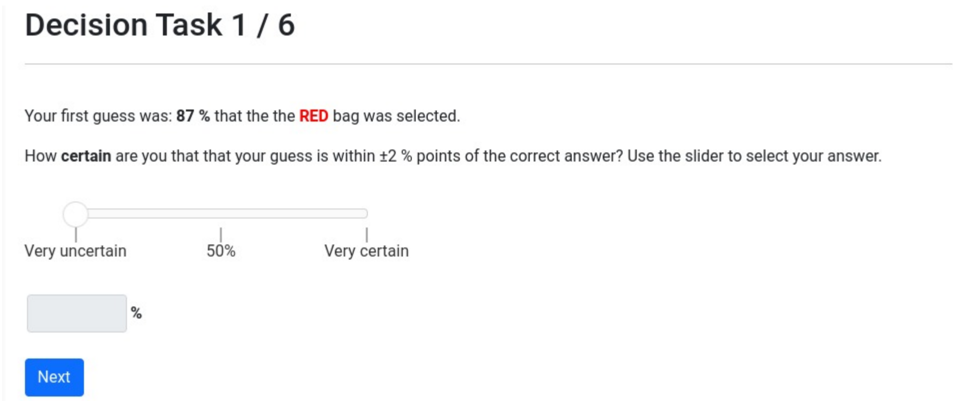


Figure D.15: Decision Tasks - certainty elicitation

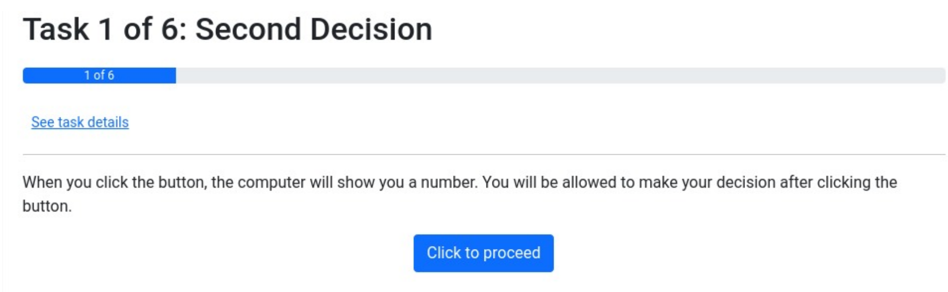


Figure D.16: Decision Tasks - Second decision

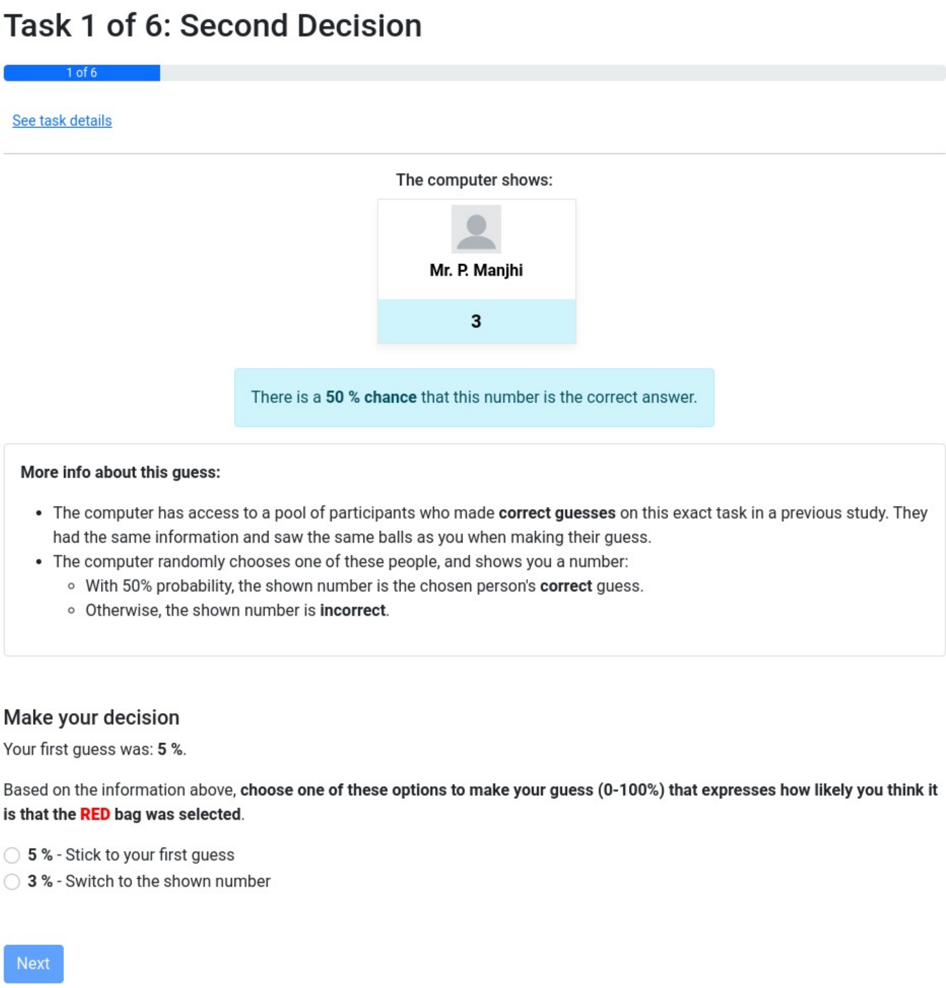


Figure D.17: Decision Tasks - Second decision

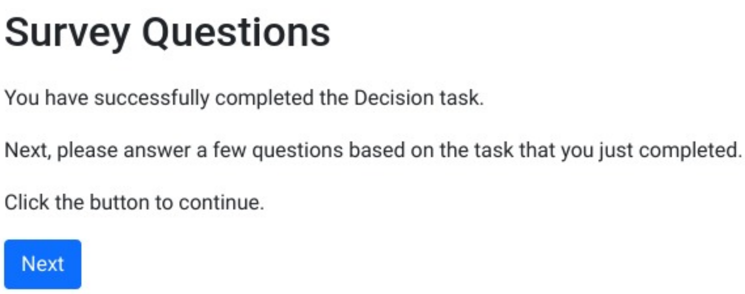


Figure D.18: Survey section

Task Reflection

Please tell us what you thought about when making the second decisions in the tasks. How did you think about the choice of using your own decision or the shown number?

Please provide a detailed response (about 20 words).

Word count:

Next

Figure D.19: Reflection question

Survey Questions

Please answer the following questions (all questions are mandatory). Accurate answers to these questions will help us in our research.

How many of your close friends have the same religion as you?

- All of them
- Most of them
- Some of them
- Hardly any of them

How many of your close friends belong to the same caste category as you?

- All of them
- Most of them
- Some of them
- Hardly any of them

Are you in favour of caste-based reservations?

- Strongly in favor
- Somewhat in favor
- Somewhat against
- Strongly against

Next

Figure D.20: Survey questions

Survey Questions

Now, you will be shown 2 questions where you have to guess the performance of people on the Decision task.

You can earn a bonus reward from these questions: One of the questions will be randomly selected for a bonus. If your answer to this question was correct, you will earn a $0.50 reward.

A correct answer to this question is an answer that is **within ±5 %** of the actual value. The actual value will be calculated from the responses provided by participants in a previous study. You will be shown the bonus amount for this part on the last page of this survey.

Click the button to continue.

Next

Figure D.21: Belief elicitation

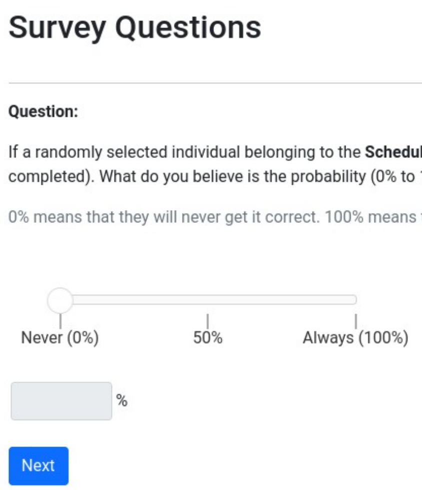


Figure D.22: Belief elicitation

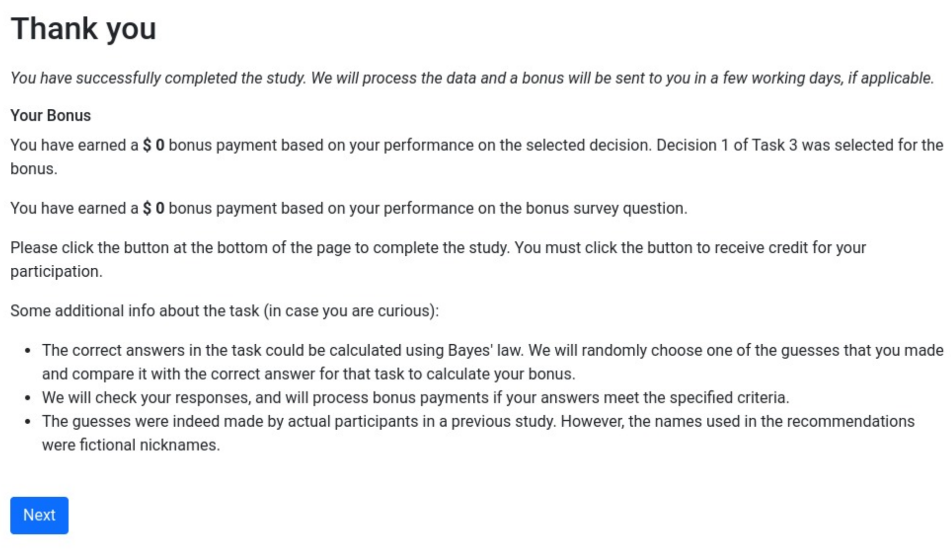
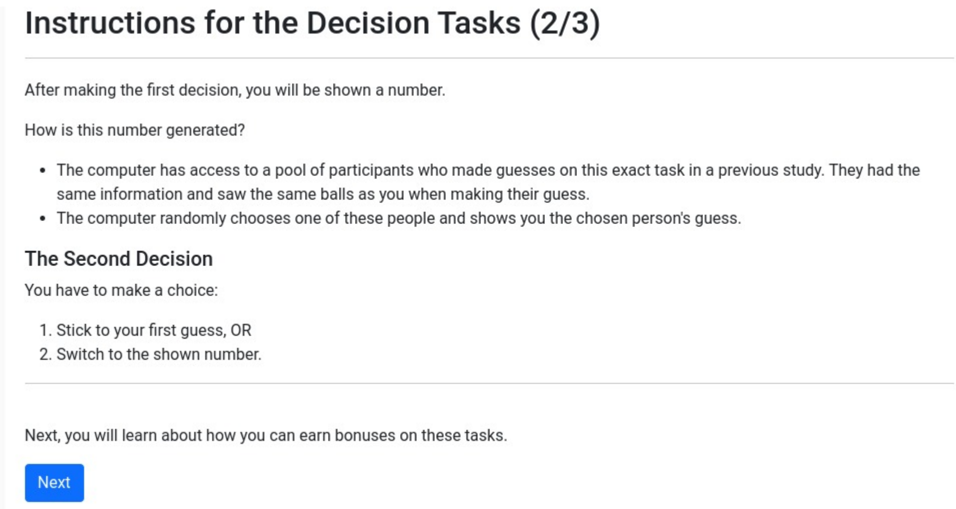


Figure D.23: Thank you and debriefing page

Caste – No Quality



Instructions for the Decision Tasks (2/3)

After making the first decision, you will be shown a number.

How is this number generated?

- The computer has access to a pool of participants who made guesses on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess.
- The computer randomly chooses one of these people and shows you the chosen person's guess.

The Second Decision

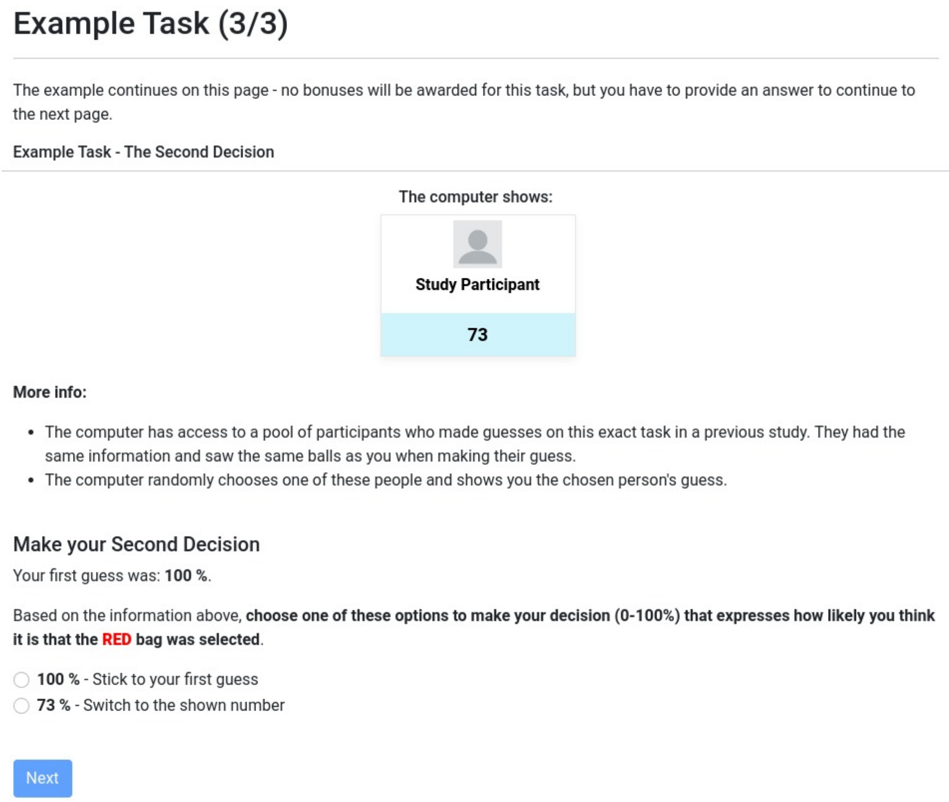
You have to make a choice:

1. Stick to your first guess, OR
2. Switch to the shown number.

Next, you will learn about how you can earn bonuses on these tasks.

Next

Figure D.24: Instructions 2/3



Example Task (3/3)

The example continues on this page - no bonuses will be awarded for this task, but you have to provide an answer to continue to the next page.

Example Task - The Second Decision

The computer shows:

Study Participant

73

More info:

- The computer has access to a pool of participants who made guesses on this exact task in a previous study. They had the same information and saw the same balls as you when making their guess.
- The computer randomly chooses one of these people and shows you the chosen person's guess.

Make your Second Decision

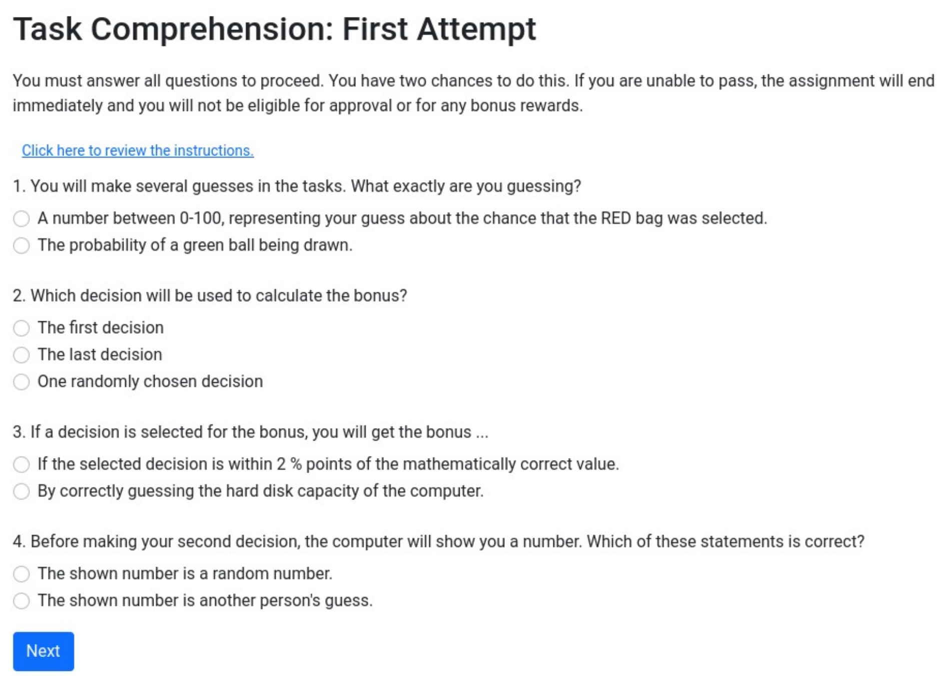
Your first guess was: **100 %**.

Based on the information above, **choose one of these options to make your decision (0-100%) that expresses how likely you think it is that the RED bag was selected**.

- ◯ **100 %** - Stick to your first guess
- ◯ **73 %** - Switch to the shown number

Next

Figure D.25: Example 3/3



Task Comprehension: First Attempt

You must answer all questions to proceed. You have two chances to do this. If you are unable to pass, the assignment will end immediately and you will not be eligible for approval or for any bonus rewards.

Click here to review the instructions.

1. You will make several guesses in the tasks. What exactly are you guessing?

- ◯ A number between 0-100, representing your guess about the chance that the RED bag was selected.
- ◯ The probability of a green ball being drawn.

2. Which decision will be used to calculate the bonus?

- ◯ The first decision
- ◯ The last decision
- ◯ One randomly chosen decision

3. If a decision is selected for the bonus, you will get the bonus ...

- ◯ If the selected decision is within 2 % points of the mathematically correct value.
- ◯ By correctly guessing the hard disk capacity of the computer.

4. Before making your second decision, the computer will show you a number. Which of these statements is correct?

- ◯ The shown number is a random number.
- ◯ The shown number is another person's guess.

Next

Figure D.26: Comprehension - No Quality

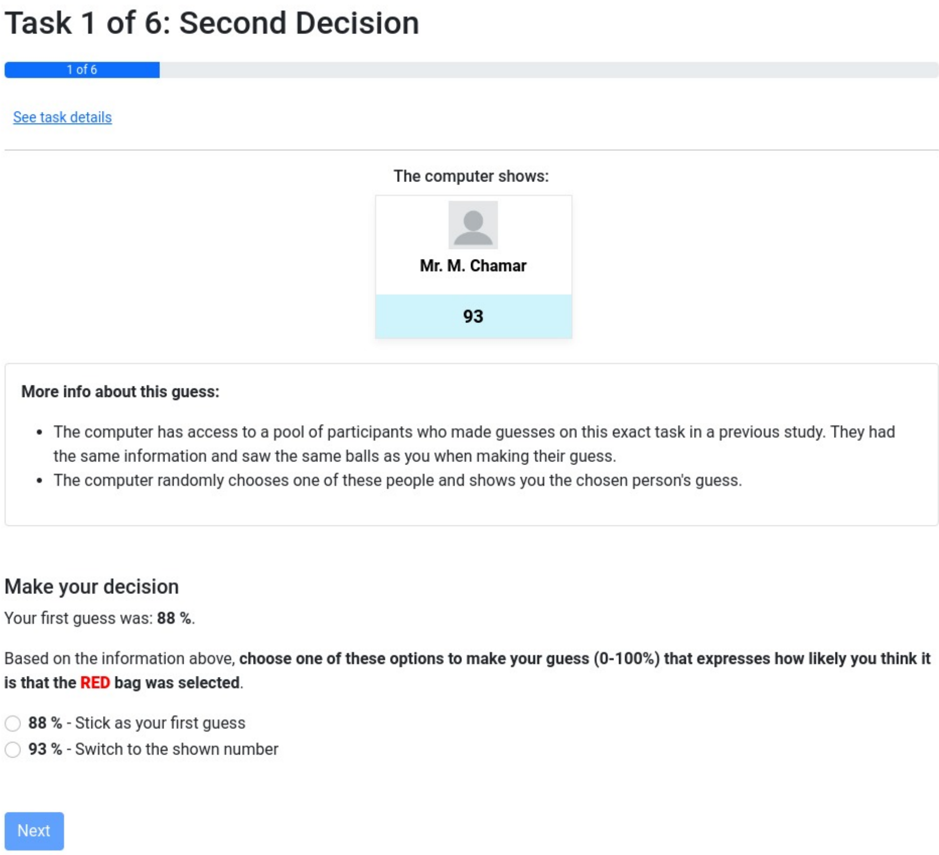


Figure D.27: Decision Tasks in *No Quality* - Second decision

Differences in other experiments

The text used to communicate the quality of the source's estimate was different in experiment *Religion, Minimal, Choice,* and *Computer*. The following screenshots are the instruction screens used in those experiments.

Instructions for the Decision Tasks (4/4)

The example continues on this page - no bonuses will be awarded for this task, but you have to provide an answer to continue to the next page. Below, you will see how to make the second decision.

The second decision

Once you submit your first decision, you will see a page similar to the one below. Here you will see another person's guess on the **same** task. This guess is drawn from a pool of guesses on this task, made by participants in a previous study.

This guess will be randomly chosen by a computer program. There is a chance with which the selected guess will be within ± 2% of the correct answer. You then have to make a choice - your second decision can be:

1. The same as your first guess, OR
2. The same as the guess made by someone else.

Remember: Each decision you make is equally likely to be chosen for the bonus, so choose the value that you think is correct.

To proceed with making your decision, click the button below. You will see a guess made by someone else on the same task.

Click to proceed

Figure D.28: Decision Tasks in *Religion* - Instructions

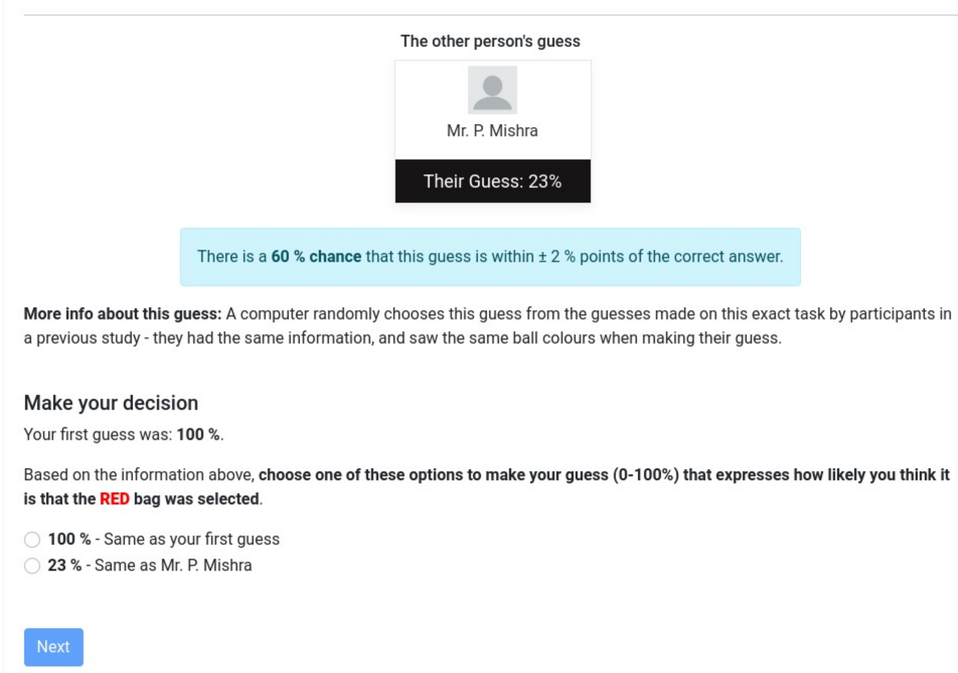


Figure D.29: Decision Tasks in *Religion* - Second guess

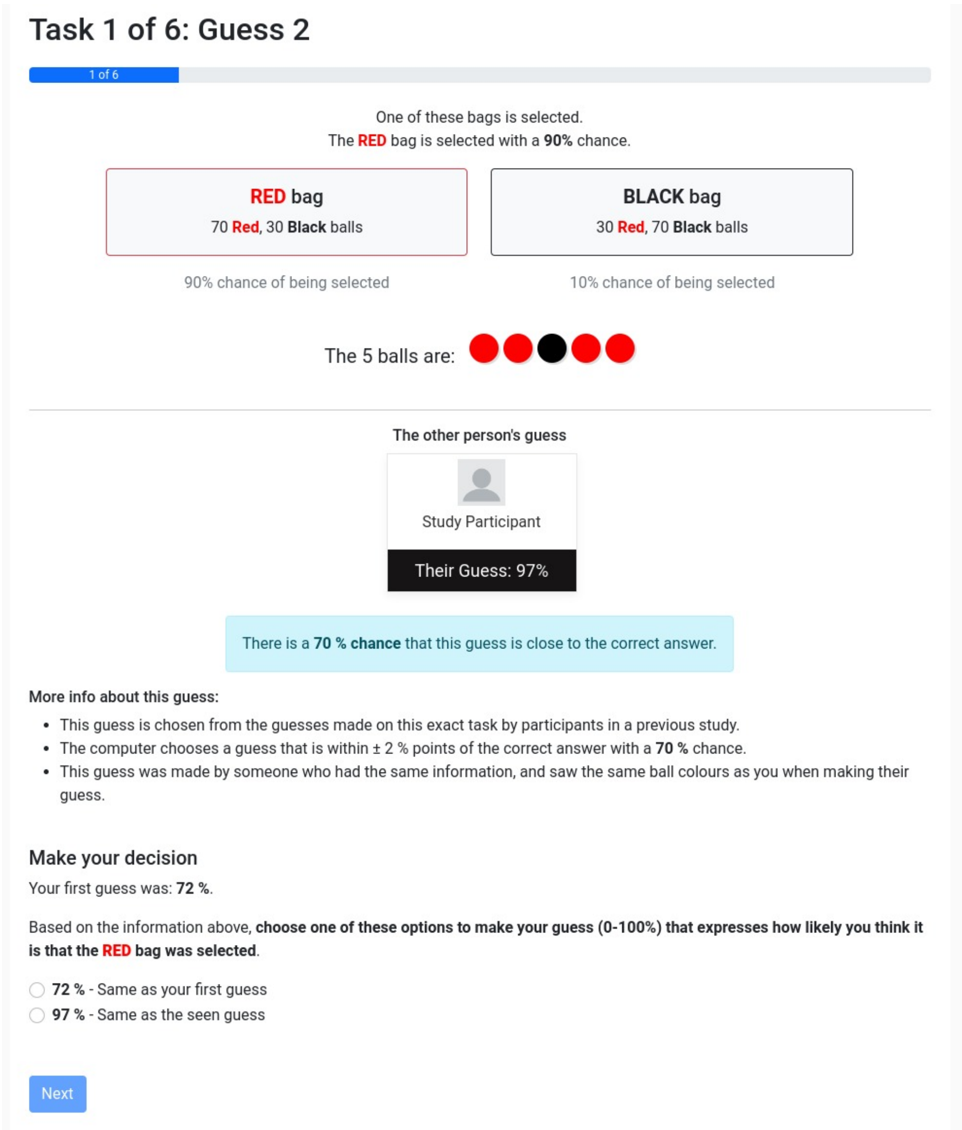


Figure D.30: Decision Tasks in *Choice* - Second guess

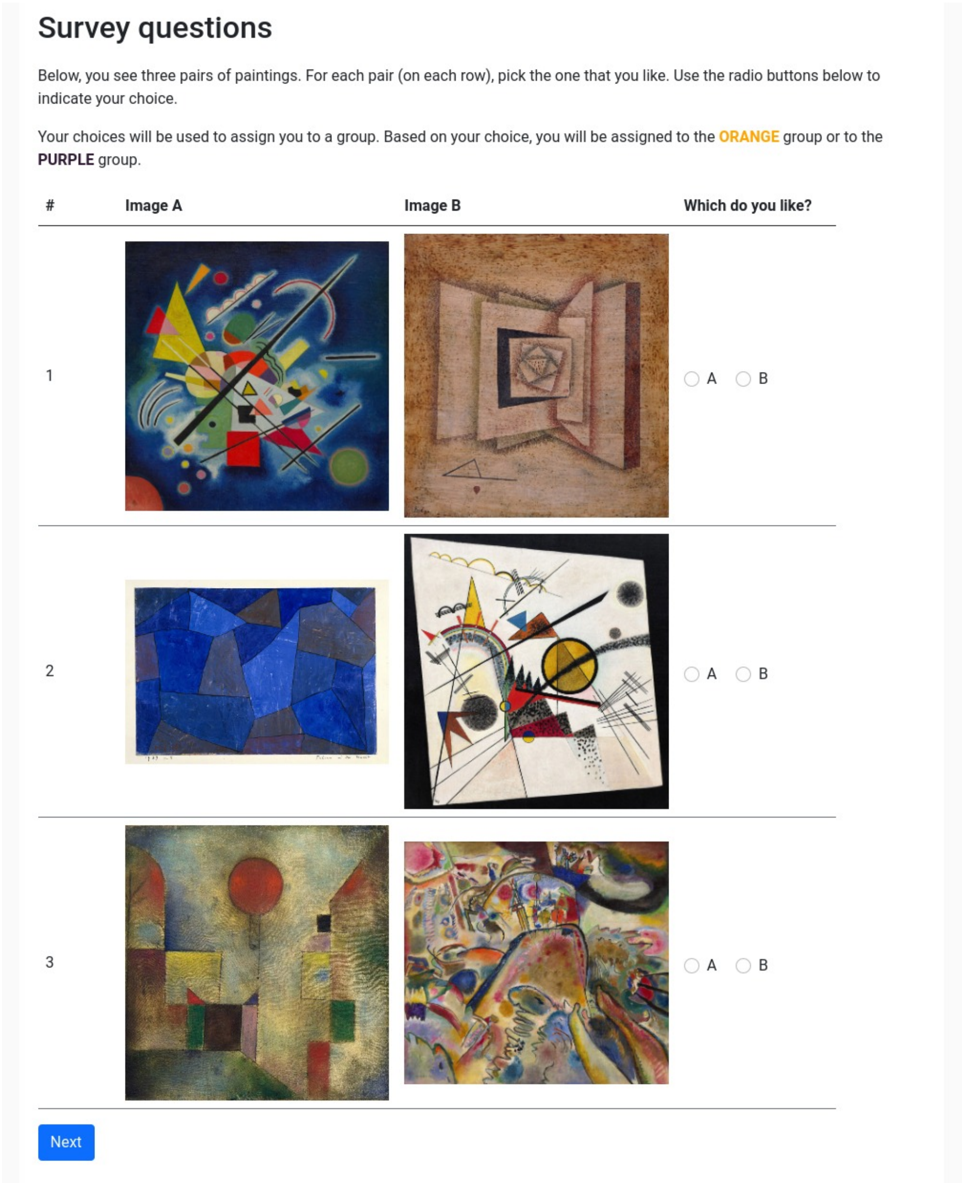


Figure D.31: Decision Tasks in *Minimal* - Klee/Kandinsky task

Instructions for the Decision Tasks (4/4)

The example continues on this page - no bonuses will be awarded for this task, but you have to provide an answer to continue to the next page. Below, you will see how to make the second decision.

The second decision

Once you submit your first decision, you will see a page similar to the one below. Here you will see a value that is chosen by a computer program in the following way: The computer chooses a value that is within ± 2 % points of the correct answer with some probability. Otherwise, the computer chooses a random number between 0 and 100. You then have to make a choice. You can choose your second decision to be:

1. The same as your first guess, OR
2. The same as the value generated by the computer.

Remember: Each decision you make is equally likely to be chosen for the bonus, so choose the value that you think is correct.

Computer

Answer: 70%

There is a **60 % chance** that this value is within ± 2 % points of the correct answer.

More info: The computer chooses a value that is within ± 2 % points of the correct answer with a **60 %** chance. Otherwise, the computer chooses a random number between 0 and 100.

Make your Second Decision

Your first guess was: **100 %**.

Based on the information above, **choose one of these options to make your decision (0-100%) that expresses how likely you think it is that the RED bag was selected**.

- **100 %** - Same as your first guess
- **70 %** - Same as the computer

Next, there will be a comprehension test to check whether you have understood these instructions clearly.

Next

Figure D.32: Instructions in *Computer* - Second guess

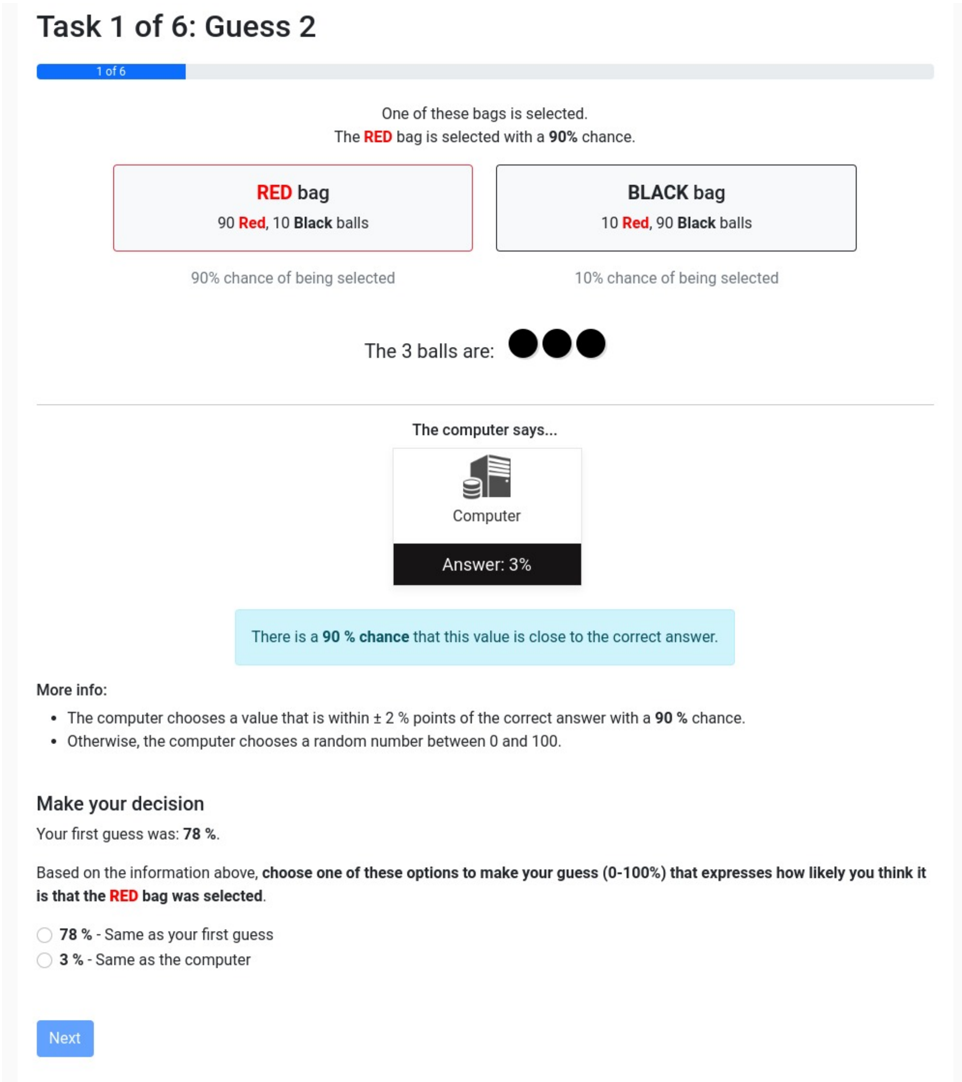


Figure D.33: Decision Tasks in *Computer* - Second guess