ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 327

The Role of Interpersonal Uncertainty in Prosocial Behavior

Anujit Chakraborty

Luca Henkel

July 2024

www.econtribute.de

UNIVERSITÄT BONN

Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

The Role of Interpersonal Uncertainty in Prosocial Behavior

Anujit Chakraborty Luca Henkel

June 25, 2024

Click here for the most recent version

Abstract

In prosocial decisions, decision-makers are inherently uncertain about how their decisions impact others' utility – we call this interpersonal uncertainty. We show that people's response to interpersonal uncertainty shapes well-known patterns of prosocial behavior. First, using standard social allocation decisions, we replicate the classic patterns of ingroup favoritism, merit-based fairness ideals, and self-favoring behavior in dictator games. We then show that these patterns also arise in non-social decisions which have no consequences for others and instead solely reflect responses to interpersonal uncertainty. Behavior across social and non-social decisions is highly correlated, and self-reported interpersonal uncertainty predicts behavior in both situations. Moreover, exogenously varying interpersonal uncertainty shifts prosocial behavior in the direction that avoids such uncertainty. Our results quantify how beliefs in the form of interpersonal uncertainty influence prosocial behavior, which we estimate to be of similar importance to social preferences.

Keywords: prosocial behavior, social preferences, ingroup versus outgroup decisions, dictator games, fairness preferences, interpersonal uncertainty
JEL Classification: C91, D01, D91

Contact: Anujit Chakraborty; University of California, Davis, chakraborty@ucdavis.edu. Luca Henkel; University of Chicago and University of CEMA, CESifo, JILAEE, IZA, luhenkel@gmail.com.
Acknowledgements: We are grateful for valuable discussions and feedback to Peter Andre, Björn Bartling, Douglas Bernheim, Felix Chopra, Bikramaditya Datta, Josh Dean, Benjamin Enke, Armin Falk, Simon Gächter, Yucheng Liang, John A. List, Ted O'Donoghue, Ryan Oprea, Devin Pope, Chris Roth, George Wu and Florian Zimmermann. Dasha Khamaza provided excellent research assistance.
Funding: Funded by the German Research Foundation (DFG) under Germany's Excellence Strategy - EXC 2126/1- 390838866 and UC Davis Small Research grants (FS24ECH).
Ethics approval: The study was approved by the ethical committees of the University of Chicago (#24-0099) and the University of California, Davis (#2134317-1).
Research transparency: All studies were preregistered at aspredicted.org (#159530, #161634, #174673, #159768, #162610, #180119). See Appendix H for details. The instructions of all studies are available in Appendix I.

1 Introduction

We as humans can only experience our own utility but not other's utility. Thus, prosocial decisions are inherently decisions made under uncertainty, where we are uncertain about how our decisions impact others' utility. We label this type of uncertainty present in prosocial decisions *interpersonal uncertainty*. Yet, economic theories of prosocial behavior abstract from such uncertainty, and empirical studies generally interpret and estimate prosocial behavior assuming certainty. However, if people respond to interpersonal uncertainty as they respond to other types of uncertainties, their response will influence their prosocial decisions.

In this paper, we show theoretically and provide experimental evidence that people's beliefs about and response to interpersonal uncertainty shape behaviors across three key paradigms of the social preference literature. Specifically, their beliefs and response reinforce ingroup favoritism in ingroup versus outgroup allocation decisions, self-favoring behavior in dictator games, and redistributive behavior when endowments are earned compared to received by windfall. Thus, observed prosocial behavior does not solely reflect social preferences. Instead, it reflects a combination of social preferences and beliefs in the form of interpersonal uncertainty, and we demonstrate how to disentangle the two. Our novel mechanism highlights the importance of beliefs in driving prosocial behavior, predicts existing evidence on the malleability of prosocial behavior, and informs interventions and policies aimed at changing prosocial behavior (e.g., promoting intergroup contact).

In our preregistered experiments, subjects make a choice in a *social* and a *non-social* decision scenario for each of the three aforementioned prosocial decision-scenarios. The *social* decision replicates a standard decision task used in the literature to elicit the respective preference/behavioral pattern. The *non-social* decision mimics the *social* decision but removes the scope for social preferences while holding the degree of interpersonal uncertainty fixed. Behavior in *non-social* then solely reflects responses to interpersonal uncertainty, allowing us to assess their relevance in generating patterns of prosocial behavior.

We illustrate our approach with the ingroup versus outgroup paradigm. In the *social* decision, a decision maker (DM) has to allocate money between two randomly matched individuals. The two individuals receive the allocated money in the form of gift cards, the decision thus has consequences for both. One of the individuals belongs to the DM's social group, making them an ingroup member, while the other is an outgroup member. Allocating more money to the ingroup member is typically interpreted as an expression of an explicit *preference* or *taste* for the ingroup. For

instance, a DM may get a higher marginal utility from the benefit received by the ingroup member (U_{in}) compared to the outgroup member (U_{out}).

We design the *non-social* decisions to rule out any such preference or taste-based channel but retain the interpersonal uncertainty. As before, the DM splits gift card money between an ingroup and an outgroup member, but now without any consequences for either. Instead, the DMs themselves are paid the sum of their matched ingroup and outgroup members' utilities U_{in} and U_{out} from receiving the gift card money. To do so, we approximate U_{in} and U_{out} by eliciting the ingroup and outgroup members' willingness-to-pay (WTP) to receive gift card money. The DMs are paid the sum of the two WTPs, weighted by DMs' allocations to the respective members. Importantly, because both WTPs contribute symmetrically to the DMs' payments, DMs no longer have any preference or taste-based reason to favor either. However, since DMs do not know the WTPs, they face interpersonal uncertainty.

In particular, people might hold different subjective beliefs about the distribution of U_{in} and U_{out} in the population. We hypothesize that interpersonal uncertainty contributes to ingroup favoritism because the DM perceives higher interpersonal uncertainty about U_{out} than U_{in} (mean-preserving spread), for instance, due to lower familiarity and fewer interactions with the outgroup. As we derive in our theoretical framework, higher uncertainty about the outgroup is sufficient to generate ingroup favoritism under risk aversion. Intuitively, allocations to the ingroup are a "safer bet" because they are less uncertain, and thus preferred by a risk-averse decision-maker.

We find that behavior in the *non-social* decisions is similar to the *social* decisions. Using shared hobbies/interests, political views, and religious beliefs as groups, subjects allocate on average 61% of the endowment to the ingroup member in the *non-social* decisions compared to 63% in the *social*. Not only the average allocations but also the distributions are similar, as we fail to reject the null of different distributions between *non-social* and *social* for shared hobbies/interests and religious beliefs. Only for political views do the distributions significantly differ. Moreover, since each subject makes both *non-social* and *social* decisions in a randomized order, we can compare behavior within-subject. Both decisions are highly correlated on the individual level ($r = 0.53$) and the median subject makes the same choice in the *non-social* and *social* decision. The presence of interpersonal uncertainty in the *non-social* decisions is thus sufficient to replicate the ingroup-favoritism of the *social* decisions.

To ensure that the observed similarity between the *social* and *non-social* decisions is not confounded by subjects being confused or inattentive about the incentives, we run three robustness experiments. In these, we systematically vary the *non-social* incentives. In the first, we increase the multiplier on the ingroup member's WTP, which

incentivizes more ingroup giving. In the second, we increase the outgroup member's WTP multiplier, incentivizing more outgroup giving. In the third, subjects are instead paid the minimum of the ingroup and outgroup members' allocation-weighted WTP, which incentivizes the 50-50 allocation. We find that behavior changes based on the incentives in the predicted directions: ingroup favoritism increases in the first treatment, flips to outgroup favoritism in the second, and favoritism vanishes in the third. These results demonstrate that behavior in the *non-social* decisions is a conscious and deliberate reaction to interpersonal uncertainty.

Our next treatments directly measure and provide causal evidence for the mechanism suggested in our framework: people perceive higher interpersonal uncertainty about outgroup members and are risk-averse towards this uncertainty. To investigate the first, we measure perceived interpersonal uncertainty using Likert scales. Subjects separately state how certain they are about ingroup and outgroup members' WTPs. Indeed, subjects perceive significantly higher uncertainty about the outgroup for all three social groups. At the same time, they perceive no difference in average WTPs. Moreover, higher relative uncertainty about outgroup members' WTPs significantly predicts stronger ingroup favoritism in both *social* and *non-social* decisions. We replicate these results in a real-effort paradigm where instead of WTP for a gift card we measure subjects' valuation for money by eliciting their willingness-to-work on real effort tasks for bonus payments.

To reveal DMs' attitudes toward interpersonal uncertainty and uncover the causal effect of their attitude on ingroup favoritism, we design a diagnostic treatment where subjects allocate money between individuals randomly chosen from two "synthetic" groups. We provide subjects with the actual valuation distribution within each group and exogenously vary this distribution across decisions. Independently, we also vary whether the group members share the DM's social group (ingroup status). The factorial variation of group information and interpersonal uncertainty allows us to isolate the marginal influence of each factor.

In the absence of ingroup/outgroup information, subjects allocate on average 60% of their endowment to the group having a lower variance in valuations, revealing an aversion to interpersonal uncertainty. When we provide both ingroup information and the distribution of valuations, subjects on average allocate 64% to the ingroup when the outgroup's valuations are more uncertain and 49% when the ingroup's valuations are more uncertain. Lastly, when all members across both groups have the same valuation, and thus interpersonal uncertainty is absent, subjects allocate 57% to the ingroup.

To quantify the influence of preferences and uncertainty responses, we regress

allocations on the treatment variations. The marginal effect of ingroup preference is similar in magnitude to the marginal response to interpersonal uncertainty, with both increasing allocations by around 7% of the endowment. We also estimate a structural model that quantifies the effect of each factor in isolation: aversion towards interpersonal uncertainty is best fit through a CRRA parameter of 0.37 and the strength of pure ingroup preference through a 7% higher allocation to the ingroup. Finally, we use the within-subject structure of our experiment to estimate behavioral types. According to our classification, 33% of subjects respond to interpersonal uncertainty but display no group preferences. In addition, 31% respond to uncertainty and display group preferences, while 20% show neither. Lastly, 17% of subjects do not respond to uncertainty but display a group preference. Taken together, we find that the majority of subjects respond to interpersonal uncertainty, and their response is quantitatively important in driving observed ingroup favoritism.

We then investigate the importance of interpersonal uncertainty for self versus other behavior using two treatments consisting of a *Self social* and a *Self non-social* decision scenario. The *Self social* decision is a standard dictator game where subjects allocate gift card money between themselves and another randomly matched person. Subjects giving more to themselves is then usually interpreted as a preference for the self. In *Self non-social*, subjects' decisions again have no consequence for others, instead, their incentives are to maximize the sum of their own WTP and the other person's WTP. Since both WTPs again contribute equally to DMs payments, a self-preference no longer predicts more allocation to the self. However, since subjects only know their own WTP but not others, interpersonal uncertainty is sufficient to generate "selfish looking choices".

We find that behavior in the *Self non-social* decision resembles behavior in *Self social*. Compared to the 69% of the endowment that subjects allocate to themselves in *Self social*, they allocate 64% to themselves in *Self non-social*, and we cannot reject equality of distributions. The similarity extends to within-subject comparisons, as the two decisions are highly correlated ($r = 0.71$), making the *Self non-social* decision one of the strongest predictors of dictator game behavior in the literature.

Our first two applications demonstrate how DMs respond when one recipient's (their own or an ingroup member's) valuation-distribution has lower interpersonal uncertainty than the other's. Our third application considers how DMs react when the recipient's valuation-distribution shifts to the right (mean-shifted distribution). In particular, we compare dictators who are allocating money from recipient's earned endowment to dictators who are allocating money from a windfall endowment. We hypothesize that dictators believe that recipient's value-distribution in the first case is

a mean-shifted version of the second case, and this explains DMs increased hesitancy to take money for themselves in the first case.

We test this hypothesis by designing the *Taking social* decision. In this modified dictator game, instead of the dictator splitting a windfall endowment as in *Self social*, the other person earned the endowment, from which the dictator can take for themselves. This difference has a significant impact on allocations. Subjects in *Taking social* allocate (take) only 41% for themselves, a significant decrease compared to the *Self social* decision. Such behavior has a natural explanation based on merit-based fairness concerns or norms: taking someone's earned money is considered more unfair than keeping money originating from a windfall.

Our *non-social* decision strips the choice of such considerations while retaining its interpersonal uncertainty. Compared to *Self non-social*, in the *Taking non-social* decision, the DM's incentive no longer depends on the recipient's WTP but on their willingness-to-accept (WTA), which measures their willingness to give up a gift card they earned. Accordingly, DMs split money to maximize the sum of their own WTP and the other individual's WTA. If DMs believe that WTA is higher, which we validate empirically, then the induced incentive leads dictators to take less money for themselves. Indeed, subjects allocate 55% to themselves in *Taking non-social*, a significant decrease compared to *Self non-social*. Moreover, allocation choices in *Taking non-social* are significantly associated with taking behavior in *Taking social*. These results suggest that the change in behavior from the dictator game to the taking paradigm is not exclusively driven by fairness considerations but instead is also influenced by the changing utilitarian calculus made under uncertainty.

Related literature. We provide evidence that interpersonal uncertainty influences patterns of prosocial behavior which have been documented across three different strands of the literature. First, a large literature has documented that people generally behave more prosocially towards ingroup members, a finding that is robust across different groups, domains, and methods (such behavior has been labeled ingroup favoritism, parochial altruism or moral universalism, see Charness and Chen, 2020; Shayo, 2020; Enke, 2023, for recent overviews).[1] Second, many studies have documented that in self versus other decisions, most people behave prosocially but tend to make choices that favor themselves more than others (see Fehr and Charness, 2023; Capraro, Halpern, and Perc, 2024, for recent overviews). Third, in allocation decisions, it has been shown that the source of the endowment matters (see Cappe-

[1] See Iyengar et al. (2019) and Böhm, Rusch, and Baron (2020) for a review of the recent literature on ingroup favoritism in political science and psychology, respectively.

len, Falch, and Tungodden, 2020, for an overview). In particular, people redistribute less if the money was earned (merited) rather than attained by windfall.[2] We propose a unified belief-based mechanism that shapes each of these patterns of prosocial behavior.[3] Our results show that due to the inherent presence of uncertainty, observed prosocial behavior cannot be interpreted solely as expressions of social preferences even in the standard elicitation tasks. In particular, our results imply that these tasks overestimate the extent of ingroup preferences (or taste-based discrimination), underestimate the degree of altruism in dictator games, and overestimate merit-based fairness preferences. We provide a methodology to separately identify and quantify the roles of beliefs and preferences in driving prosocial behavior.

With our subjective uncertainty-based explanation of prosocial behavior, we relate to a recent literature that explains a range of behavioral patterns through people's cognitive response to (subjective) uncertainty. Enke and Graeber (2023) investigate how people's uncertainty over the optimal decision influences choice under risk, belief formation, and forecasts. In the domain of intertemporal decisions, a series of theoretical studies show that risk and time preferences closely intertwine when DMs are uncertain about future consumption (Sozou, 1998; Dasgupta and Maskin, 2005; Halevy, 2008; Chakraborty, Halevy, and Saito, 2020) or preferences (Amador, Werning, and Angeletos, 2006; Chakraborty, 2021). While this literature focuses on characterizing a logical equivalence between subjective uncertainty and risk or time preference patterns, we study the connection between subjective uncertainty and prosocial behavior. In particular, our non-social treatments allow us to assess the extent to which subjective uncertainty in the form of interpersonal uncertainty drives standard patterns of social behavior.[4]

To isolate the importance of interpersonal uncertainty in prosocial choices, we construct the non-social decisions by stripping the original social decisions of all motivations based on social preferences. Similarly, Oprea (2024) and Enke, Graeber, and Oprea (2023) construct diagnostic decisions by stripping risk or discounting-based motivations from standard risky and intertemporal tasks to isolate the role of complexity on decision-making under risk and time.

[2]See Ruffle (1998), Cherry (2001), Cherry, Frykblom, and Shogren (2002), Cherry and Shogren (2008), Oxoby and Spraggon (2008), and Krupka and Weber (2013).

[3]Previous belief-based explanations of prosocial behavior have been mainly applied to strategic interactions, such as trust or reciprocity (Berg, Dickhaut, and McCabe, 1995; Fehr and Gächter, 2000) and intentions (Falk, Fehr, and Fischbacher, 2008).

[4]We thus differ from papers investigating prosocial behavior under experimenter-induced objective risk over consequences to study ex-post versus ex-ante fairness (e.g., Brock, Lange, and Ozbay, 2013), the use of risk to act selfishly (Exley, 2016), or the relation between risk preferences and giving under risk (Cettolin, Riedl, and Tran, 2017). Further, Cappelen et al. (2022) and Cappelen, De Haan, and Tungodden (2024) study redistribution decisions when the source of inequality is uncertain.

2 Conceptual framework

Our central premise is that when choosing between different actions of which at least one has consequences for others, a DM perceives *Interpersonal Uncertainty* about how those consequences impact others' utility. That is, she is uncertain about the utility that others receive from the outcome created by her actions.[5] Specifically, we analyze a DM who has to allocate $\$100$ between two recipients. In such an allocation decision, interpersonal uncertainty can be understood as the subjective uncertainty the DM perceives about how much utility or valuation each recipient derives from each allocated dollar. We will show that a simple model of interpersonal uncertainty can generate canonical patterns from the literature on prosocial behavior.

2.1 Assumptions about interpersonal uncertainty

For simplicity, we assume that the DM is probabilistically sophisticated and believes that dollars are valued non-negatively. Interpersonal uncertainty then means that the DM believes the per-dollar valuation of recipient j is distributed as $v_j \sim f_j$, where f_j is a probability distribution with non-negative support contained in $[0, b]$, and F_j is the corresponding CDF. We define $S_j(x) = \int_0^x F(y)$.

Next, we assume that these belief distributions have two key features. First, a DM understands that different recipients might derive different values from the same allocated dollar amount based on their personalities, past experiences, socioeconomic status, or tastes. Thus, the belief distributions over the valuations of others are non-degenerate.

Second, the belief distributions for different recipients systematically differ, depending on the DM's familiarity with the recipients, or the source of the $\$100$ endowment. For instance, suppose one recipient shares their hobbies/ religious/ political interests with the decision-maker (ingroup member) while the other does not (outgroup member). This makes the allocation decision an ingroup versus outgroup tradeoff, which will be our leading example. Facing this tradeoff, a DM might think that shared interests or identity with a recipient is indicative of shared past experiences, economic status, and tastes. As a consequence, DMs may feel less familiarity and thus perceive higher interpersonal uncertainty about the outgroup. Similarly, in situations involving the DM herself as one of the two recipients, DMs may naturally face higher uncertainty about others than about themselves since we are most familiar with our own tastes and circumstances. In other situations, subjects might

[5] Therefore, it differs from uncertainty about the mapping between actions and outcomes, for instance the uncertainty whether a donation will actually be delivered to a recipient.

think that one recipient is systematically more likely to have higher valuations than another recipient. Formally,

Definition 1. *Subjects perceive a higher interpersonal uncertainty for recipient 2 than recipient 1 if* $S_1(x) \leq S_2(x)$ *for all* x *and* $S_1(y) < S_2(y)$ *for some* y*. Subjects perceive a mean-shifted interpersonal uncertainty for recipient 2a compared to recipient 2b, if there exists* $c \in \mathbb{R}_{++}$ *such that for all* x*,* $F_{2a}(x + c) = F_{2b}(x)$*.*

The condition for higher interpersonal uncertainty is best understood as a generalization of "f_2 *is a mean-preserving spread of* f_1*"* or equivalently "f_2 *is second-order stochastically dominated by* f_2*"*, because the latter notions are defined identically with the additional condition that f_1 and f_2 have equal means. We will use the concept of higher interpersonal uncertainty to characterize the optimal allocation x^*. In comparison, we will use the mean-shift concept to understand how x^* changes when the DM's beliefs about a particular recipient's (say 2's) valuation-distribution shifts to the right (from 2b to 2a).

2.2 Choice behavior under interpersonal uncertainty

We investigate the case of unbiased utilitarian preferences, which means the utility the DM receives from allocating $x \in [0, 100]$ to the ingroup and $(100 - x)$ to the outgroup is $u_{UTIL} = v_1 x + v_2(100 - x)$. As v_1, v_2 are random variables, she maximizes expected utility over the potential utilitarian outcomes:

$$EU(x) = E_{v_1 \sim f_1, v_2 \sim f_2} U\left(v_1 x + v_2(100 - x)\right) \tag{1}$$

where $U' > 0$ and $E_{v_i \sim f_i}$ is the expectation with respect to f_i.

Given this setup, the optimal allocation depends crucially on the response to uncertainty as characterized by U, and on the belief distributions f_1 and f_2. We will generally assume that $U'' < 0$ which implies that the DM dislikes higher variance over potential utilitarian outcomes. If both f_1 and f_2 are degenerate with different expected values, the DM will allocate 100 to the recipient with the higher expected value.[6] If both distributions are non-degenerate, Theorem 1 provides the optimal solutions and serves as our prediction for both the *social* and *non-social* decisions we later employ in our experiments.

Theorem 1. *Suppose individual* i *has unbiased utilitarian preferences and is risk-averse* ($U'' < 0$)*. If* f_1 *and* f_2 *are non-degenerate, independent probability distributions, then*

[6] In the trivial case of degenerate distributions with equal expected value, the optimal allocation is non-unique, as the DM is indifferent between all possible allocations.

i) Equal division: If $v_1 \overset{d}{=} v_2$ (i.e, $f_1 = f_2$) then i's optimal allocation is $x^* = 50$.
ii) Ingroup favoritism: If f_2 is a mean preserving spread of f_1, then i's optimal allocation is $x^* \in (50, 100)$.
iii) Comparative statics over x^*: Suppose the valuations of the two groups are distributed as v_1 and $c+v_2$ for some constant c and independent random variables $v_1 \sim f_1, v_2 \sim f_2$. Under arbitrary CARA preferences[7], or under CRRA coefficient< 1, the optimal allocation satisfies $dx^*/dc \leq 0$.

For the proof, see Appendix Section A. Part (i) follows from symmetry: a risk-averse DM hedges against interpersonal uncertainty by allocating equally among ex-ante symmetric recipients. (ii) shows that if the DM perceives a higher interpersonal uncertainty about one of the recipients, she allocates more to the other recipient. Accordingly, a DM who perceives higher interpersonal uncertainty about the outgroup member will allocate more money to the ingroup, even if they believe that on average, ingroup and outgroup members benefit equally from receiving money. Similarly, a DM who perceives higher interpersonal uncertainty about other's utility than their own utility will allocate more money to themselves, even if they care about others equally and think that on average, everyone benefits equally from money. This motivates our experiments studying the ingroup versus outgroup paradigm and the self versus other paradigm in Sections 3 and 4. Note that in (ii) we use the assumption of equal expected values simply as a benchmark: our key insight is that interpersonal uncertainty can generate ingroup favoritism despite equal expected values.

Finally, part (iii) shows that the DM would decrease the allocation to the ingroup (or the allocation to herself in the dictator game) if her belief about the outgroup's valuation mean-shifts to the right.[8] For example, if a DM perceives mean-shifted interpersonal uncertainty when allocating a recipient's earned money compared to allocating windfall money (thus, perceiving higher c in the former case), then she would keep less for herself (lower x^*) in the former case. This motivates our experiments studying the giving versus taking paradigm in Section 5.

Will every commonly used welfare criterion deliver the results of Theorem 1 under the right parameters given our assumptions about interpersonal uncertainty? In Appendix C, we show that Rawlsian preferences are insensitive to higher interpersonal uncertainty. We will use this result later in a robustness analysis to show that

[7]For a utility function $U(w)$, the coefficient of absolute risk aversion (ARA) is defined as $r_1(w) = \frac{-U''}{U'}$ and relative risk aversion (RRA) is defined as $r_2(w) = \frac{-wU''}{U'}$. CARA and CRRA imply r_1 and r_2 are constant respectively.

[8]Under extreme risk aversion, when c increases, the marginal return from the states with high v_2 is so low that on the margin, subjects might prefer to allocate more to v_1 to safeguard their utility in the states where v_2 is low.

people respond to our induced incentives in the expected direction.

Relation to preferences. The economic literature generally interprets ingroup favoritism purely as an expression of ingroup preferences, modeled as a higher utility weight for ingroup compared to outgroup members (e.g., Tabellini, 2008). In psychology, it is often interpreted as an expression of moral values (e.g., Graham et al., 2013). Similarly, various explanations for the fact that people allocate more, but not all of the endowment to themselves in dictator games have been brought forward (e.g., Capraro, Halpern, and Perc, 2024). Most of these models either implicitly or explicitly assume that DMs weight their own utility differently than others' utility.

However, under interpersonal uncertainty, risk aversion is sufficient to generate ingroup favoritism or self-favoring behavior, differential utility weights are no longer necessary. Importantly, our framework does not imply the absence of social preferences. To the contrary, the described patterns of prosocial behavior emerge precisely because people have social preferences: they care about others' utility, but as this utility is unobserved, they face uncertainty.

3 Ingroup versus outgroup paradigm

We start by studying ingroup versus outgroup decisions before expanding to further prosocial decisions in later sections.

3.1 Experimental design

The experimental sessions using the ingroup-outgroup paradigm (and the other paradigms introduced later) feature two distinct decision situations: *social* decisions and *non-social* decisions. The *social* decision is a classical decision task from the literature on prosocial behavior. Thus, for the ingroup versus outgroup paradigm, the *social* decision is a "bystander" money-allocation game – one of the standard experimental decision tasks used to identify differential attitudes towards ingroup and outgroup members (e.g., Chen and Li, 2009; Enke, Rodríguez-Padilla, and Zimmermann, 2022).[9] The game features three individuals, (i) a decision-maker (DM), (ii) one individual who shares a social group with the DM (ingroup member), and (iii) another individual who is a member of a different group than the DM (outgroup member). The DM is asked to allocate a fixed amount of money between the ingroup and outgroup

[9]The first bystander allocation game involving ingroup and outgroup members was conducted by Tajfel et al. (1971).

members. The degree to which DMs allocate more money to the ingroup member reveals their degree of ingroup favoritism.[10]

Our novel contribution is to design and implement a novel decision situation, the *non-social* decisions. In these decisions, we remove any altruistic motivations but retain the inherent interpersonal uncertainty about the utility others experience. Next, we explain the details of the *social* decisions and the *non-social* decisions.

***Ingroup social* decisions.** In total, decision-makers face three *Ingroup social* decisions, in each allocating \$100 between one ingroup and one outgroup member. Specifically, they allocate money (i) between someone who "shares your interests/hobbies" versus "has different interests/hobbies than you", (ii) between someone who "shares your political views (e.g., a fellow left-winger, or a fellow right-winger, etc.)" versus someone who "has different political views than you" and (iii) between someone who "shares your religious beliefs (e.g., a fellow Christian, or a fellow atheist, etc.)" versus someone who "has different religious beliefs than you".[11] The allocated money is sent to the ingroup and outgroup member six weeks from the date of the experiment in the form of Amazon gift card money. Thus, DM's allocation decisions have consequences for the utility of the individuals involved.

***Ingroup non-social* decisions.** In our *Ingroup non-social* decisions, we remove any consequences the decision has to other individuals. Instead, DM's choice solely determines their own payoff. DMs split \$100 between an ingroup and outgroup member, using the same groups as in *social*, and the DM's payoff Π is determined by the following formula:

$$\Pi\left(x_{in}, x_{out}\right) = x_{in} \cdot WTP_{in}/100 + x_{out} \cdot WTP_{out}/100.$$

where x_{in} is the money split in favor of the ingroup member, and $x_{out} = 100 - x_{in}$ is the money split in favor of the outgroup member. WTP_{in} and WTP_{out} denote the ingroup and outgroup member's respective WTP for a \$100 Amazon gift card to be received in six weeks, elicited using a valuation task (explained below). To scale the incentive, the WTP is divided by 100, representing an individual's WTP per gift card dollar. For example, if the DM split \$40 and \$60 in favor of the ingroup and outgroup member respectively, and the elicited WTP of "\$100 Amazon gift card

[10]Particularly, ingroup favoritism is identified independent of the decision-maker's self-interest. Past research has shown that behavior in such bystander allocation games shows a high test-retest correlation, works equally well when posed hypothetically and incentivized, and is highly correlated with related psychological questionnaires (Enke, Rodríguez-Padilla, and Zimmermann, 2022).

[11]The wording is taken from Enke, Rodríguez-Padilla, and Zimmermann (2022).

money received in 6 weeks" for the ingroup member were \$80 and for the outgroup member \$60, then the DM's payoff would be

$$\Pi\left(40, 60\right) = 40 \cdot 80/100 + 60 \cdot 60/100 = 68.$$

By using the WTP of ingroup and outgroup members, we induce utilitarian preferences because we incentivize the DMs to maximize the sum of the WTPs, weighted by the allocations made in their favor. Since DMs do not know the actual WTPs of the matched individuals, this interpersonal uncertainty transforms the *social* decision into an uncertain subjective lottery choice. At the same time, because the WTP is elicited over the same object that is distributed in the *Ingroup social* decision, we keep the degree of interpersonal uncertainty constant between the *Ingroup social* and *Ingroup non-social* decision. Importantly, the ingroup and outgroup member's WTP enter the utilitarian payoff function symmetrically, so any differences in allocations are driven by differences in uncertainty about the WTPs. We can thus use the comparison of the *Ingroup social* and *Ingroup non-social* decision to assess the relevance of interpersonal uncertainty in driving ingroup favoritism.

Valuation task. To elicit the willingness-to-pay (WTP), we use a standard multiple-price-list (MPL). Subjects face a series of binary decisions between (i) receiving a \$100 Amazon gift card in six weeks and (ii) a monetary amount paid today which increased across decisions.[12] This procedure reveals the current-day dollar equivalent of receiving gift card money.[13]

Minimizing inattention and confusion. A principal concern when interpreting behavior in the *non-social* decisions is that subjects are inattentive to the incentive structure or misunderstand the parameters of the decision. We employ several measures to mitigate the scope for these confounding factors. First, before completing the *non-social* decisions, decision-makers complete the valuation task themselves. That is, they face the WTP elicitation, which familiarizes them with the calculation of the WTP for the incentive. Second, we included several comprehension questions that test whether DM's understood that the *non-social* decisions only have consequences for themselves, not for the other individuals. If they did not answer all questions correctly, we explained them their errors and highlighted the correct answers. This

[12]We enforced single switching by automatically filling the list above and below subjects' choices.

[13]We used Amazon gift card money because it is easy to anonymously pay online, the specific gift cards are non-refundable and non-fungible, and because subjects' valuation generally differs from the dollar value of the gift card. We implemented the time lag to generate additional variation in subjects' valuation.

procedure makes it particularly salient that the *non-social* decisions are different from the *social*. Third, to further minimize inattention, we include an explicit disclaimer on the *non-social* decision screens that states "Reminder: your choice only determines your own payment, it does not affect the two individuals." On the decision screen, we also provided DM's with the option to revisit the instructions.

In addition to these measures to mitigate the role of inattention and confusion, we designed a series of robustness treatments to assess the extent to which limited attention or confusion could drive behavior in our experiments. These treatments vary elements of the incentives, and are described in detail in Section 3.3.

Procedure. We randomized the order of decisions. Half of the decision-makers first face the *social* decision and then the valuation task and *non-social* decision. The other half first face the valuation task and *non-social* decision, and then subsequently the *social* decision. We did not announce beforehand that other decisions would follow the initial decisions, therefore minimizing the scope for contagion from one treatment to the other. This design allows us to analyze within-subject behavior, and compare behavior between-subject by only looking at the first set of decisions.

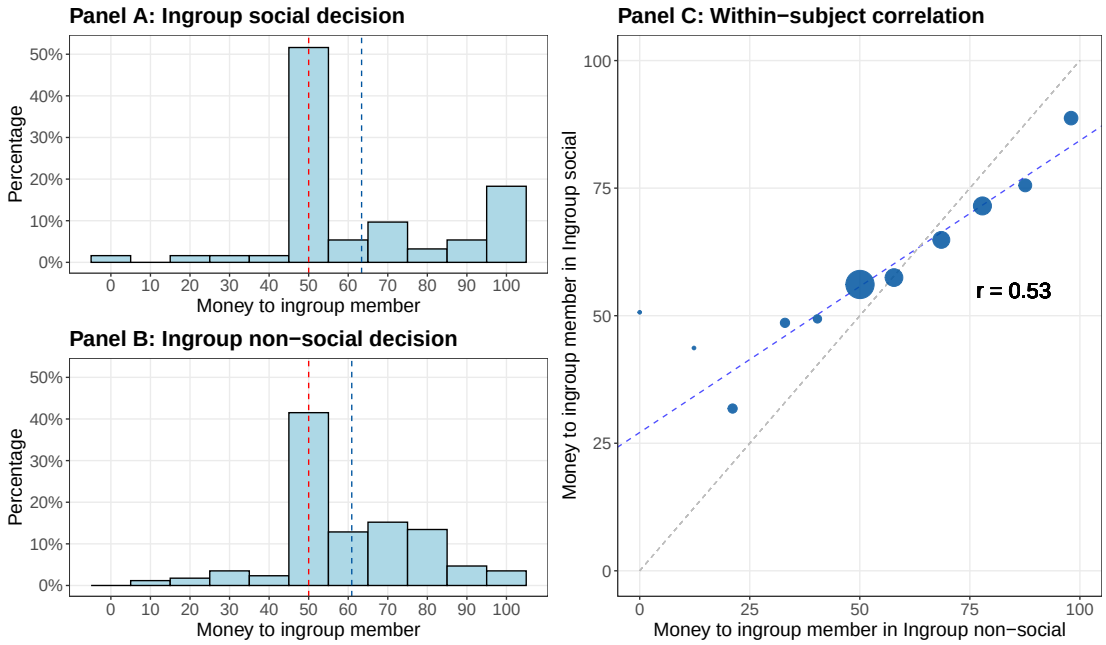
Data. In total, 119 subjects participated in the ingroup experiment, with 62 subjects first facing the *Ingroup social* decision and 57 subjects first facing the *Ingroup non-social* decision. For this and all further experiments, we used Prolific to recruit online participants living in the US. We chose Prolific due to its status as one of the leading market research companies used in social science research and because their participants have been shown to provide high-quality responses in terms of comprehension and attention (Eyal et al., 2021; Gupta, Rigotti, and Wilson, 2021). All experiments were preregistered, see Appendix H for details. We used oTree (Chen, Schonger, and Wickens, 2016) for programming the graphical user interface. Subjects spent a median of 10 to 12 minutes in the experiments and received as compensation the equivalent of an hourly wage between $10 and $12 per hour. In each experiment, one randomly selected subject out of the participating subjects had one randomly selected decision implemented with real consequences.

3.2 Results

We start with the between-subject comparison of the *social* and *non-social* decisions before moving to the within-subject comparison.

***Ingroup social* decisions.** In the *Ingroup social* decisions, subjects allocate on average $57.48, $71.05, and $61.61 to the ingroup member when they share the same

Figure 1: Main results ingroup versus outgroup decisions



Notes: **Panel A and B:** Histogram of *Ingroup social* (Panel A) and *Ingroup non-social* (Panel B) decisions. The x-axis denotes the amount of gift card money (out of \$100) that subjects allocate to the ingroup member instead of the outgroup member. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Ingroup social* (Panel A), the decisions have consequences for the ingroup and outgroup members. In *Ingroup non-social* (Panel B), the decisions have consequences only for the subjects, with their payoff depending on the ingroup and outgroup member's WTP for the gift card. **Panel C:** Binned scatter plot of *Ingroup social* and *Ingroup non-social* decisions. The blue dotted line displays the linear fit of a regression of the *Ingroup social* on *Ingroup non-social* decisions. The correlation coefficient is $r = 0.53$. For all three panels, the binwidth is 10. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs), displaying $n = 186$ decisions by 62 subjects in Panel A, $n = 171$ decisions by 57 subjects in Panel B, and $n = 357$ decision-pairs by 119 subjects in Panel C.

interests/hobbies, the same political views, or the same religious beliefs, respectively. In all three cases, we can reject the hypothesis of no ingroup-favoritism ($p < 0.01$, one-sample Wilcoxon tests). Figure 1 panel A displays the distribution pooled over the three decisions, which replicates the typical distributional pattern found in the literature (e.g., Enke, Rodríguez-Padilla, and Zimmermann, 2022). In 46% of the decisions, subjects display ingroup-favoritism by allocating strictly more than 50% to the ingroup. Outgroup-favoritism is found in 8% of decisions, and in the remaining 46%, subjects allocate 50/50. In total, 73% of subjects display ingroup-favoritism in at least one decision, making it the prevalent mode of decision-making.

***Ingroup non-social* decisions.** Importantly, a similar pattern emerges in the *Ingroup non-social* decisions. Here, subjects allocate on average \$56.86, \$65.02, and \$60.81 when splitting in favor of ingroup members sharing the same interests/hobbies,

same political views, and same religious beliefs. As before, we find significant ingroup-favoritism in all three cases ($p < 0.01$, one-sample Wilcoxon tests), even though the decisions have no consequences for ingroup or outgroup members. See Panel B of Figure 1 for the distribution of the pooled decisions. In 61% of the *Ingroup non-social* decisions, subjects display ingroup-favoritism by allocating strictly more than 50% to the ingroup. Outgroup-favoritism is found in 11% of decisions, and in the remaining 28%, subjects allocate 50/50.

Comparing *Ingroup social* and *non-social*. We cannot reject that average ingroup allocations are equal between *Ingroup social* and *non-social* decisions in any of the three cases ($p = 0.59$ for hobbies/interests, $p = 0.22$ for political views, $p = 0.38$ for religious beliefs, unpaired Wilcoxon tests). Further, we cannot reject that the allocation distributions are equal when the groups concern hobbies/interests and religious beliefs ($p = 0.15$ and $p = 0.10$, Kolmogorov-Smirnov test). We can only reject the null of equal distributions in the case of political views ($p = 0.004$, Kolmogorov-Smirnov test). Thus, our *non-social* setup where decisions have no consequences for either group member closely replicates ingroup versus outgroup attitudes from the standard social setup. The most notable difference between the two decisions is the extent to which subjects display maximal ingroup favoritism, i.e., give the entire endowment to the ingroup member. Almost 20% of decisions in *social* display this pattern, mostly stemming from the decisions involving political view groups. In contrast, less than 5% of *non-social* display maximal ingroup favoritism.

Within-subject comparison. Next, we compare behavior between *Ingroup social* and *non-social* on the individual level by including also the second set of decisions of each subject. We replicate the previously reported between-subjects results also within-subject, see Appendix D.1. Importantly, we find no evidence for order effects, supporting the validity of our within-results (see Appendix E for details). This allows us to correlate behavior in *Ingroup social* with *non-social*. Panel C of Figure 1 displays the distribution of each individual social and non-social decision pair in a binscatter-plot. As the figure shows, the two are highly related: ingroup favoritism in *Ingroup non-social* predicts ingroup favoritism in *Ingroup social*, with a correlation coefficient of $r = 0.53$. Therefore, the same subjects that display ingroup favoritism when decisions have consequences for others also display it when their decisions solely affect their own payoff, with the payoff depending on other's WTPs.

Result 1. *We find ingroup-favoritism in Ingroup non-social, which retains interpersonal uncertainty but removes any consequences for ingroup or outgroup members. The*

distribution of behavior is similar to Ingroup social, which features consequences, and decisions in the two situations are strongly correlated on the individual level.

3.3 Robustness

Our main results show a high degree of similarity between the *Ingroup social* and *Ingroup non-social* decisions. Next, we present a series of robustness treatments to establish that this similarity is not driven by subjects being confused or inattentive to the experimental design.

3.3.1 Subjects understand and react to utilitarian incentives

If subjects are confused or do not pay attention to the *non-social* incentives, they may treat the *non-social* decisions as *social* decisions. This, in turn, would artificially increase the similarity between the two types of decisions. To test for these confounds, we designed two variations of the *Ingroup non-social* decisions, the *Ingroup incentive* and *Outgroup incentive* treatments.

Design. In the *Outgroup incentive* and *Ingroup incentive* treatments, we vary the utilitarian incentives by changing the weights that are put on the ingroup and outgroup members' WTPs. All other aspects of the *Ingroup non-social* decisions are left unchanged. In *Outgroup incentive*, we increase the weight on the outgroup member's WTP to be three times as high as the ingroup member's WTP. A subject's payoff thus becomes:

$$\Pi\left(x_{in}, x_{out}\right) = x_{in} \cdot WTP_{in}/100 + 3 \cdot x_{out} \cdot WTP_{out}/100$$

Similarly, in *Ingroup incentive* we increase the weight on the ingroup member's WTP to be three times as high as the outgroup member's WTP:

$$\Pi\left(x_{in}, x_{out}\right) = 3 \cdot x_{in} \cdot WTP_{in}/100 + x_{out} \cdot WTP_{out}/100$$

If subjects respond to the incentives we induce in the *non-social* decisions, ingroup favoritism should decrease in *Outgroup incentive* and increase in *Ingroup incentive*. In total, 120 subjects participated in this robustness experiment, facing both treatments in random order. As before, we start with the between-subject comparison by focusing on behavior in the first-assigned set of choices.

Results. We find that subjects respond to changes in the induced utilitarian incentives. Compared to an average giving of \$60.89 to the ingroup member across the three social groups in the baseline *Ingroup non-social* of section 3.2, subjects give on average \$42.16 to the ingroup member in *Outgroup incentive* and \$67.98 in *Ingroup incentive*, a significant difference in both instances ($p < 0.001$, unpaired Wilcoxon test). In particular, behavior switches to outgroup favoritism in *Outgroup incentive*, as the average is significantly smaller than the even split ($p < 0.001$, one-sample Wilcoxon test). Regarding the distributions, the fraction of subjects displaying outgroup favoritism is 54% in *Outgroup incentive* and 10% in *Ingroup incentive*, while the fractions displaying ingroup favoritism are 33% and 72%, respectively. See Appendix Figure B.1 for the corresponding histograms. The significant shifts in the distributions ($p < 0.01$, Kolmogorov-Smirnov test) show that the changes in the averages are not driven by a small minority of subjects reacting strongly, but a substantial fraction of subjects.

Indeed, the within-subject analysis reveals that subjects change their behavior in 81% of decisions following the incentive change between the two treatments.[14] Accordingly, in 19% of decisions are subjects unresponsive to changes in the *non-social* incentives, indicating inattention or confusion. Comparing the behavior in these situations to behavior in *Ingroup social* of the main experiment, average ingroup giving is slightly less (\$60.84 compared to \$63.38), and also the fraction of choices favoring the ingroup is lower (34% of decisions compared to 46%) while 50/50 splits are more frequent (60% to 46%).

These results provide evidence against limited attention or confusion driving our results of the previous section. Changing a single number in the incentive formula reverses the direction of favoritism from ingroup to outgroup favoritism. Moreover, in those cases where choices indicate inattention or confusion about the incentives, ingroup favoritism is, if anything, less prevalent relative to the main experiment.

3.3.2 Inducing Rawlsian preferences changes behavior

The previous treatments show that inducing utilitarian incentives induce ingroup favoritism. But would inducing any class of preferences lead to a replicating of the ingroup favoritism found in *social* also in the *non-social* or is utilitarianism special in this regard? Suppose we induce Rawlsian preferences, one of the most important alternatives to utilitarianism. Would subjects still choose identically to the utilitarian case, showing ingroup favoritism as before? Or would their choices adapt to the

[14] We replicate the between-subject results in the within-subject case, see Appendix Section D.2.

new incentives? To answer these questions, we designed the *Non-social minimum* treatment, which induces Rawlsian preferences.

Design. In the treatment *Non-social minimum*, subjects face *non-social* decisions, but instead of incentivizing a utilitarian preference, we incentivize a Rawlsian welfare function. Specifically, a subject's payoff is calculated as:

$$\Pi\left(x_{in}, x_{out}\right) = \min\{x_{in} \cdot WTP_{in}/100, x_{out} \cdot WTP_{out}/100\}$$

Thus, we incentivize them to choose the allocation that maximizes the utility of the worse-off recipient, irrespective of group affiliation. All other aspects of the decisions are identical to the *Ingroup non-social* decisions. In total, 62 subjects participated in the *Non-social minimum* treatment.

The *Non-social minimum* treatment also helps us test the following confound: DMs may be inattentive to the change in instructions between *Ingroup social* and *Ingroup non-social*. Hence *Ingroup non-social* might mechanically replicate *Ingroup social*. DMs only become attentive once the differential incentives induced in *Ingroup incentive* and *Outgroup incentive* introduce a payoff-asymmetry. This form of selective limited attention could in principle explain the similarity between *social* and *non-social* decisions as well as the response to the incentive treatments. The *Non-social minimum* treatment puts this to the test, because the Rawlsian payoff rule treats ingroup and outgroup symmetrically yet incentivizes a different allocation. As we show in Section 2, if the WTP distribution for the outgroup is a mean-preserving spread of the ingroup's WTP and subjects are risk-averse, the optimal choice under utilitarian preferences is $100 > x_{in} > 50$, implying ingroup favoritism. In contrast, as we show in Appendix C, the optimal choice under Rawlsian preferences is $x_{in} = x_{out} = 50$, implying no favoritism in either direction.

Importantly, we kept the decision screen in *Non-social minimum* identical to the screen in *Ingroup non-social*. Hence, if subjects are inattentive or confused about the *non-social* incentives so that they erroneously think they face the *Ingroup social* choice instead, we should observe ingroup favoritism in *Non-social minimum*. In contrast, if they are attentive and understand the incentive, we should observe no favoritism in *Non-social minimum*, neither in the ingroup direction nor the outgroup direction.

Results. As predicted, we find that ingroup favoritism is eliminated under the Rawlsian incentive. On average, subjects allocate \$51.31 to the individual sharing their interests/hobbies, \$52.85 to the individual sharing political views and \$49.77 to the individual sharing religious beliefs. Accordingly, the treatment did not only signif-

icantly reduce ingroup favoritism (in all three cases $p < 0.01$, unpaired Wilcoxon tests), but eliminated it altogether, as we can no longer reject that average ingroup giving is different from the 50/50 split ($p = 0.31$, $p = 0.13$ and $p = 0.95$ respectively, one-sample Wilcoxon tests). For the distributions of decisions, see Appendix Figure B.2, which further show that subjects respond strongly to the induced incentives in the expected direction: the percentage of decisions that implement exactly a 50/50 split increases from 32% in *ingroup non-social* to 58% *Ingroup non-social minimum*, and a further fraction of 22% subjects is within \$10 of the even split. Similarly, the percentage of highly unequal allocations, as measured by giving less than \$30 or more than \$70 to the ingroup member, drops from 32% to 12%.

3.3.3 Results replicate when changing the valuation tasks determining the incentives

As final robustness check, we assess the degree to which our results depend on the use of gift card money. In the previous experiments, we used gift card money as part of the valuation task for the incentives in the *non-social* decisions and as the good that is allocated in the *social* decisions. In the *Ingroup effort* robustness treatment, we remove the gift cards.

Specifically, in the *Ingroup effort social* decision, subjects allocate a plain amount of \$10 between an ingroup and an outgroup member. In *Ingroup effort non-social*, instead of paying subjects based on the sum of the ingroup and outgroup members' WTP for a \$100 gift card money, subjects are paid based on the members' willingness-to-work (WTW) for a \$10 bonus payment. Eliciting the WTW for the ingroup and outgroup member works as follows: they face a series of binary decisions between (i) completing a number of real effort tasks to receive the \$10 bonus payment and (ii) not working. We present the decision in a multiple-price-list format, with the number of completed tasks required for the bonus ranging from 0 to 30 in increments of 2. We then define the WTW as the number of tasks for which subjects first switch from preferring to work to not working. The task itself is a simple slider task, adopted from (Gill and Prowse, 2012, 2019). Each individual task consists of moving 30 sliders to the middle position within a 60-second time limit. A task counts as completed if 90% or more sliders are moved to the middle position.

For this robustness exercise, we also change the implementation probability. One might worry that the previous method of implementing the decision of one subject per experiment induces an excessively small winning probability. Therefore, in *Ingroup effort social* and *Ingroup effort non-social*, one out of every ten subjects had

one of their decisions implemented. All other elements of the *social* and *non-social* decisions were held constant. In total, 121 subjects participated in this robustness experiment.

Results. The robustness treatments closely replicate our main findings. In the *Ingroup effort social* decisions, subjects allocate on average $6.27 to the ingroup member, i.e., 63% of the endowment, almost identical to the 61% found in *Ingroup social*. Ingroup favoritism is found in 41% of decisions (46% in *Ingroup social*), outgroup favoritism in 2% (8%) of decisions, and the remaining 57% (46%) of decisions are 50/50 allocations. In *Ingroup effort non-social*, the average giving to the ingroup member is 62% of the endowment compared to 61% in *Ingroup non-social*, with 57% (61%) of decisions showing ingroup-favoritism, 12% (11%) showing outgroup-favoritism, and the remaining 31% (28%) decisions are 50/50 allocations. As before, we also find a strong within-subject correlation of 0.68 between *Ingroup effort social* and *Ingroup effort non-social*. Appendix Figure B.3 displays the corresponding between-subject histograms and the within-subject bin-scatter plot.

3.4 Interpersonal uncertainty as mechanism

According to our theoretical framework of Section 2, even if the ingroup and outgroup are perceived as having equal valuations on average, interpersonal uncertainty generates ingroup favoritism through a combination of two factors: first, subjects perceive higher interpersonal uncertainty for the outgroup than for the ingroup members. Second, subjects are averse to higher uncertainty. Accordingly, we first measure the perceived difference in interpersonal uncertainty and correlate it with choice behavior. In a second step, we measure the causal effect of exogenously varied interpersonal uncertainty on choice behavior.

3.4.1 Perceived interpersonal uncertainty differs between groups and predicts behavior

Elicitation. To elicit perceived interpersonal uncertainty, we asked *Ingroup social* DMs the following question after the *social* decisions, separately about their ingroup and outgroup members: "How certain are you about how much the individual (. . .) would value Amazon gift card money?" In the *Ingroup effort social* treatment (Section 3.3.3), "Amazon gift card money" was replaced by "the bonus payment". Subjects could respond on an 11-point Likert scale from *Very uncertain* to *Very certain*. For the analysis, we re-code the variable so that higher values indicate higher uncertainty.

We then create a relative uncertainty measure at the subject level by subtracting every subject's reported uncertainty over an ingroup member's valuation from their reported uncertainty over an outgroup member's valuation.

Results. We find that subjects indeed perceive higher outgroup uncertainty: In *Ingroup social* (*Ingroup effort social*, respectively) the self-reported uncertainty is on average 0.57 (0.95) Likert scale points higher under the ingroup-outgroup classification for different hobbies/interests, 1.01 (1.39) higher for different political views, and 0.45 (1.01) higher for religious beliefs. All differences are significantly different from zero ($p < 0.05$, paired Wilcoxon tests). Pooled across the three groups, subjects perceive 0.68 (1.12) higher outgroup uncertainty. With respect to the distribution of differences, in 31% (34%) of cases subjects report higher uncertainty for the outgroup; in 12% (5%) they report higher uncertainty for the ingroup and in the remaining 58% (61%) cases, subjects report no difference.

Importantly, differences in uncertainty predict choices: higher uncertainty about the outgroup is associated with stronger ingroup favoritism in *Ingroup social* ($r = 0.30$, $p < 0.001$), *Ingroup effort social* ($r = 0.43$, $p < 0.001$), as well as in *Ingroup non-social* ($r = 0.17$, $p = 0.002$), and *Ingroup effort non-social* ($r = 0.28$, $p < 0.001$). See Figure B.4 in the Appendix for the corresponding binned scatter plots.

Robustness. One might wonder if interpersonal uncertainty, elicited after a *social* or *non-social* choice, could be confounded by the prior choice. For example, subjects might report motivated beliefs that somehow justify their prior choice(s). Hence, in a separate robustness treatment for beliefs (*Ingroup belief measurement*, $n = 59$), we elicited subjects perceived interpersonal uncertainty without them making any *social* and *non-social* choices. We replicate our patterns from the main treatment. Subjects report 0.81 Likert scale points higher uncertainty for individuals having different hobbies/interests, 0.93 for different political views, and 0.66 for religious beliefs (all three significant at $p < 0.01$, paired Wilcoxon tests). Pooled across the three social groups, in 37% of cases subjects report strictly higher outgroup uncertainty, in 58% no difference and in 5% higher uncertainty for the ingroup.

As a validation exercise, at the end of the survey, we asked the same subjects to report their perceived interpersonal uncertainty for two artificial groups with exogenously induced objective WTP-distributions, where the second group's WTP-distribution was the mean-preserving spread of the first's. 74% (87% respectively) of subjects report strictly (weakly) higher interpersonal uncertainty about the second group, validating the sensitivity of the Likert-scale measure. See Appendix G for details.

In addition, we asked subjects to report their estimated mean WTP for both ingroup and outgroup members to investigate if mean-differences could also contribute to the observed ingroup favoritism in the *non-social* treatment. On average, subjects report ingroup members to have a gift card valuation of \$87.85, \$89.64, and \$89.14 for shared hobbies/interests, political views and religious beliefs respectively. For outgroup members, they report valuations of \$88.20, \$86.69, and \$87.85. Thus, on average, subjects' estimates do not differ between ingroup and outgroup ($p = 0.90$, $p = 0.06$, and $p = 0.37$, paired Wilcoxon tests).[15] Pooled across all three groups, subjects believe in 30% of cases that the average ingroup WTP is higher, in 49% of cases they believe them to be equal and in 21% of cases they believe the average outgroup WTP to be higher.

Taken together, these findings demonstrate that subjects indeed perceive higher uncertainty about the outgroup's valuation but do not systematically think that the ingroup has higher valuations on average (no mean shifted distribution).

3.4.2 The causal effect of interpersonal uncertainty on behavior

In our model, risk aversion towards interpersonal uncertainty generates ingroup favoritism whenever the perceived interpersonal uncertainty is higher for the outgroup. In a new treatment (*Ingroup uncertainty*) with 120 subjects, we test this channel by decoupling and independently varying interpersonal uncertainty and ingroup status.

Design. DMs faced seven decisions in random order. In each decision, DMs were endowed with \$10 to allocate between Group A and Group B. Each group consists of two individuals, whom we label as recipients. Recipients were participants of an earlier study in which they provided their willingness-to-work (WTW).[16] DMs were told that after they made their allocation decision, a randomly chosen recipient from each group would receive the money allocated to that respective group. Between decisions, we systematically varied (i) uncertainty over the recipients' WTW within each group and (ii) their ingroup/ outgroup affiliation. Thus, the seven decisions can be classified into four types:

1. *Uncertainty without group information* decision: the two recipients of Group

[15] Moreover, on average, beliefs are not systematically biased: the average estimates are close to the actual average WTP that is observed in the experiment, which is \$87.

[16] Recall that the WTW elicits the willingness to complete tasks to receive a \$10 bonus payment, with tasks ranging from 0 to 30, thus serving as a proxy for the utility value of the bonus payment, measured on an effort scale. To familiarize DMs in *Ingroup uncertainty* with the WTW, they first worked on an example slider task and subsequently faced the WTW elicitation themselves. See Section 3.3.3 for details on the task and elicitation.

A have the same WTW of 12, while the recipients of Group B have a WTW of 4 and 22, respectively. Thus, while the recipient who eventually gets the money from Group A would have a fixed WTW of 12, the recipient receiving the money from Group B could have a WTW 4 or 22. Thus, the DMs face higher interpersonal uncertainty about the WTW of Group B recipients according to Definition 1 in Section 2.

2. *Group information without uncertainty* decision (x2)[17]: the two recipients of Group A are ingroup members, while the two recipients of Group B are outgroup members. Moreover, all four recipients have the same WTW of 12.

3. *High uncertainty on ingroup members* decision (x2): the two recipients of Group A are ingroup members, one having a WTW of 4 and the other of 22. The two recipients of Group B are outgroup members, both having a WTW of 12. Hence, the ingroup has a higher WTW variation.

4. *High uncertainty on outgroup members* decision (x2): the two recipients of Group A are ingroup members, both having a WTW of 12. The two recipients of Group B are outgroup members, one having a WTW of 12 and the other of 22. Hence, the outgroup has the higher WTW variation.

The *uncertainty without group information* treatment reveals DMs' attitude towards higher uncertainty about WTWs without the confound of ingroup preferences. Because the expected value of WTWs is higher in Group B[18], DMs who still allocate more to Group A reveal their aversion to the uncertainty in WTWs, and hence their risk aversion. The *group information without uncertainty* decision, on the other hand, reveals ingroup favoritism in the absence of interpersonal uncertainty. The last two decision-situations reveal the extent to which interpersonal uncertainty influences ingroup favoritism.

Importantly, the WTW information is based on recipients' *stated* WTW, not their actual exerted effort. While facing the WTW elicitation, recipients knew that with a 50% chance they would complete their selected WTW tasks and with a 50% chance they would not work but later potentially receive money from other participants (the DMs) who observe their WTW. We then only matched the latter half of subjects with

[17]The "x2" indicates that there were two independent decisions of this type, one based on the ingroup of similar hobbies and another based on the ingroup of similar political views.

[18]Intrinsically, this argument holds when recipients' value of $10 is directly equal to their WTWs, as well as when recipients' valuation is equal to an increasing and convex disutility function of the WTW. The latter assumption is well supported by empirical evidence on real effort tasks (see e.g., Augenblick, Niederle, and Sprenger, 2015; Gill and Prowse, 2019).

the DMs of the *Ingroup uncertainty* experiment. Thus, none of the recipients had worked based on their WTW choices, a fact we saliently communicated to the DMs. This ensures that the DMs have no fairness reason to be partial to one recipient over the other based on exerted effort.

DMs faced decisions 2, 3 and 4 twice, once when the ingroup/ outgroup was based on shared/different hobbies or interests and once when the ingroup/ outgroup was based on shared/different political views.[19] When presenting the WTW and social group information for the decisions, we randomized the position of the group on the screen (left or right) and which information was presented first, to balance the presentation of the two pieces of information.

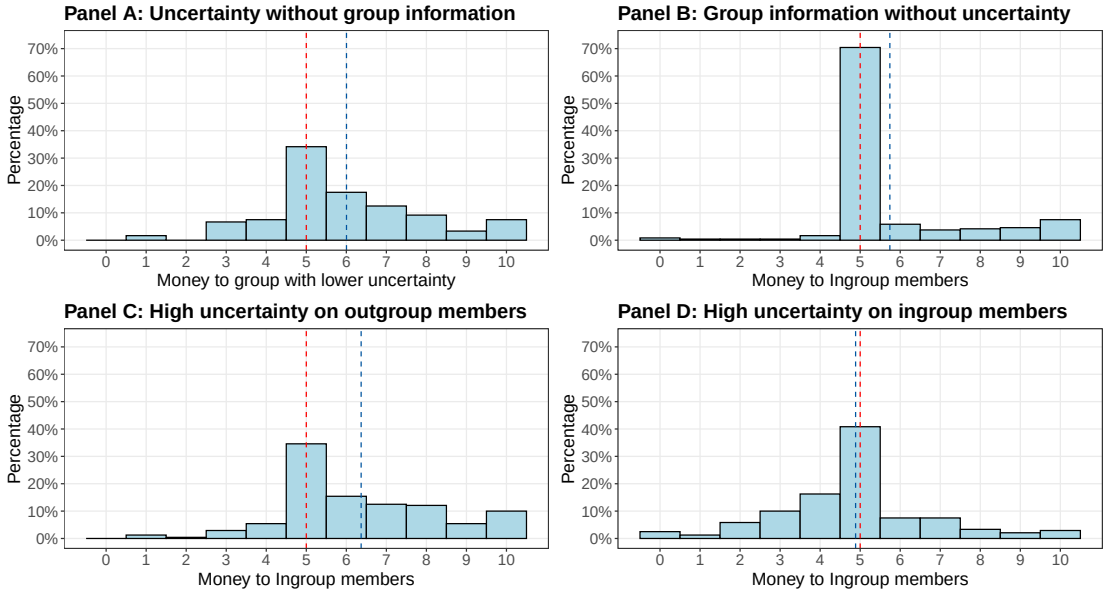
Results. Figure 2 displays the distribution of choices. Starting with Panel A, we observe that the majority of subjects are risk-averse towards interpersonal uncertainty even in the absence of group information: 54% allocate more than 50/50 to the group with lower WTW variance, 27% allocate 50/50 and a minority of 19% allocate more to the group with the higher WTW variance. On average, subjects allocate \$6.00 to the group with lower WTW variance, which is significantly more than the 50/50 benchmark ($p < 0.001$, one-sample Wilcoxon tests).

In the case of all recipients having the same WTW (Panel B), subjects allocate on average \$5.74 to the ingroup (pooling the decision across both social groups), which is significantly different from the 50/50 split ($p < 0.01$, one-sample Wilcoxon test). This quantifies the extent to which ingroup preferences drive ingroup favoritism since interpersonal uncertainty is absent in this decision.

Panels C to D then document how interpersonal uncertainty being higher in the ingroup or the outgroup changes the magnitude of ingroup favoritism. The amount allocated to the ingroup increases from \$5.74 to \$6.37 when the outgroup is more uncertain (has higher variation in WTWs), and decreases to \$4.89 when the ingroup is more uncertain (both $p < 0.001$ compared to the no uncertainty decision, paired Wilcoxon tests). Moreover, changing the uncertainty changes the entire distribution of choices. When the outgroup has the higher WTW variance, the modal (62%) DM displayed ingroup favoritism, while 12% displayed outgroup favoritism and 26% had the 50/50 split. Once the recipients of both groups have equal WTWs, the model subject (66%) now chooses the 50/50 split. Here, only a minority of 29% and 5% strictly favor the ingroup and outgroup, respectively. Lastly, when the ingroup has the higher WTW variance, the modal subject (40%) switches to outgroup favoritism,

[19]We did not include groups involving religious beliefs because based on the previous results, we expected similar behavior, and we wanted to avoid adding three more decisions.

Figure 2: Main results ingroup versus outgroup decisions



Notes: **Panel A:** The x-axis denotes the amount of money (out of $10) that subjects allocate to the group with a lower variation in their willingness-to-work (WTW), and thus, a lower interpersonal uncertainty. In **Panels B to D**, the x-axis denotes the amount of money given to the group where both recipients share a social group with the decision-making compared to the other group where both recipients have a different social group. In Panel B, all group members have the same WTW. In Panel C, the outgroup has the higher variation in WTWs, while in Panel D the ingroup has the higher variation. As such, Panel B displays the case of no interpersonal uncertainty, Panel C of higher interpersonal uncertainty in the outgroup, and Panel D the reverse. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. For all four panels, the binwidth is 10. Decisions are pooled across the two social groups involved (shared hobbies/interests and political views), displaying $n = 120$ decisions in Panel A, and $n = 240$ decisions in Panel B to D, made by 120 subjects.

while 31% of choices show ingroup favoritism and 29% are 50/50 choices. These results establish that responses to interpersonal uncertainty causally influence the extent of ingroup favoritism.

Information choice. After making the seven decisions that exogenously informed subjects which group had ingroup affiliation and/or which group had a higher WTW-variance, subjects made an eighth and final allocation decision. In this decision, they were not given any information that distinguished one group from the other ex-ante. They were only informed that the two groups were different: (i) both recipients of one group (which group was unspecified) shared their hobbies/interests (or shared their political views), while both recipients of the other group did not, and (ii) both recipients of one group (which group was unspecified) had a WTW of 12, while the two recipients of the other group had a WTW of 4 and 22. However, ex-ante, they did not know which group had the ingroup members, or which group had a lower

variance in WTW, or if ingroup members had lower/ higher WTW variance.

Subjects then could choose to learn one of the two dimensions along which the groups differed. That is, they could either learn which group contained only ingroup members and which contained only outgroup members, or, they could learn which group had WTW variation and which had not.[20] Which information subjects choose reveals which factors are of primary importance in their decision process.

We find that 81% of subjects choose to learn which group has the high WTW variation, and only 19% choose to learn the ingroup-outgroup information. Accordingly, in our setting, subjects prefer receiving information about interpersonal uncertainty over information about social group membership, indicating its relevance for decisions.

3.5 The quantitative importance of interpersonal uncertainty

The previous sections provided causal evidence that both interpersonal uncertainty and ingroup preferences independently influence ingroup favoritism. Next, we quantify the relative magnitudes of each channel using a reduced-form analysis and a structural model of prosocial decisions.

3.5.1 Reduced form analysis

The design of the *Ingroup uncertainty* experiment allows us to answer two questions that quantify the importance of each factor in isolation. First, on average, what fractions of ingroup favoritism are driven by interpersonal uncertainty versus ingroup preferences? Second, what percentage of subjects make allocation choices that reveal the relevance of interpersonal uncertainty and/or ingroup preferences?

Average ingroup favoritism. To estimate the effects of uncertainty and preferences on ingroup preferences, we use the following model: $alloc_{i,d} = c_0 + c_1 unc_d + \varepsilon_{i,d}$, in which $alloc_{i,d}$ denotes the allocation to the ingroup by individual i in decision d. We normalize the variable by subtracting 5 from the actual giving so that the 50/50 split benchmark implies $alloc_{id} = 0$. The variable unc_d indicates the presence of interpersonal uncertainty in the decision. The variable is equal to 1 for decisions with *Higher WTW uncertainty on outgroup members*, equal to -1 for decisions with *Higher WTW uncertainty on ingroup members*, and equal to 0 when uncertainty is absent for both groups. Thus, c_1 measures how much the allocation is affected by uncertainty

[20] We balanced the presentation of both pieces of information by randomizing the order in which the information was introduced and displayed for the choice.

in either direction. Accordingly, ingroup preferences are measured by the constant c_0 that captures whether subjects allocate more to the ingroup on average even in the absence of interpersonal uncertainty ($unc_d = 0$).

Column (1) of Table 1 displays the results of this specification. We find that both ingroup preferences and interpersonal uncertainty significantly influence on behavior, and their influence is *nearly equal in magnitude*.

In column (2), we estimate an alternative specification where we split the *Interpersonal uncertainty* variable into two indicators for the two uncertainty decisions. All coefficients are significant at the 0.001 level, with the influence of adding uncertainty on the ingroup being slightly larger than adding uncertainty on the outgroup.

Type analysis. We can further exploit the within-subject structure to identify distinct behavioral types at the subject level. We say subjects reveal a *group-based preference* if they choose a different allocation than 50/50 in *at least one* of the two decisions that provided *group information without uncertainty*. Accordingly, a subject does not reveal group-based preference if they allocate 50/50 in both those decisions. We say that a subject *responds to interpersonal uncertainty* if, for both social groups (interest and political views), their *outgroup uncertainty* decision was different from their *ingroup uncertainty* decision. Hence, if subjects choose the same allocation in outgroup uncertainty and ingroup uncertainty for at least one social group, they are not categorized as responding to uncertainty.[21] With this categorization, we find that 20% of subjects neither respond to uncertainty nor display a group preference. 33% of subjects respond to uncertainty but do not display a group preference, while 17% do not respond to uncertainty but display a group preference. Finally, 31% both respond to uncertainty and display a preference. Hence, interpersonal uncertainty is relevant for 64% of all subjects, while group preferences are relevant for 48%.

Information choice type analysis. We further use subjects' choice on whether they prefer to receive information about the recipients' WTWs or group affiliations to validate our type categorization. That is, we compare the fraction of subjects choosing the WTW information instead of the group affiliation information across our four behavioral types. In total, 92% of subjects who respond to uncertainty but display no group preference choose the WTW information. This fraction decreases to 78% for those who both respond to uncertainty and reveal a group preference, and decreases

[21] Thus, we use a more conservative identification criterion for the response to uncertainty, because we require subjects to respond to uncertainty across both pairs of choices. In contrast, for identifying group-based preferences only one choice needs to be different from the "no favoritism" benchmark. In total, 79% of subjects respond to at least one change in uncertainty from ingroup uncertainty to outgroup uncertainty.

further to 50% for those who reveal a group preference but do not respond to uncertainty. Accordingly, our type categorization predicts subjects' information choices in the expected direction.

Table 1: The influence of changes in interpersonal uncertainty on ingroup favoritism

	Dependent variable:	
	Allocation to ingroup	
	(1)	(2)
Constant (Ingroup preference)	0.741^{***}	0.666^{***}
	(0.132)	(0.104)
Interpersonal uncertainty	0.742^{***}	
	(0.115)	
Higher uncertainty ingroup		-0.856^{***}
		(0.144)
Higher uncertainty outgroup		0.629^{***}
		(0.135)
Subjects	120	120
Observations	720	720
R^2	0.096	0.095

Notes: The table shows OLS estimates. The dependent variable is the amount subjects allocate to the ingroup (out of \$10), subtracted by five (thus values range from -5 to 5). "Interpersonal uncertainty" is an indicator variable that is equal to 1, 0, or -1 when the decision has *High uncertainty on ingroup members*, *Group information without uncertainty*, and *High uncertainty on outgroup members* respectively. "Higher uncertainty ingroup" ("Higher uncertainty outgroup") is an indicator variable equal to 1 when the decision has *High uncertainty on ingroup members* (*High uncertainty on outgroup members*) and is zero otherwise. Standard errors (in parentheses) are clustered at the subject level. Significance levels: *p$<$0.1, **p$<$0.05 and ***p$<$0.01.

3.5.2 Structural model

A complementary approach to the previous reduced form analysis is the use of a structural model.

Setup. Suppose the representative DM has to distribute M units between two individuals 1 and 2. For each of the individuals, the group identity can take one of three values: $G(i) \in \{in, out, \emptyset\}$, where $\emptyset$ means unknown. Suppose the DM believes that the valuations of money received by the two individuals 1 and 2 are distributed as f_1 and f_2 respectively, and suppose $x_{IU}(f_1, f_2, \gamma)$ is the choice that maximizes the

expected (utilitarian) utility (see equation 1) given the two distributions. To parameterize risk aversion, we assume CRRA utility $u(w) = \frac{w^{1-\gamma}}{1-\gamma}.$, with γ as the risk aversion parameter. We assume that the DM's optimal allocation to individual 1 in observation j is as follows:

$$x_{1j} = \begin{cases} \frac{M}{2} + b + \varepsilon_j & \text{if } f_1 = f_2, G(1) = in, G(2) = out \\ x_{IU}\left(f_1, f_2, \gamma\right) + \varepsilon_j & \text{if } f_1 \neq f_2, G(1) = G(2) \\ a_{IU} \cdot x_{IU}\left(f_1, f_2, \gamma\right) + a_{ING} \cdot \left(\frac{M}{2} + b\right) + \varepsilon_j & \text{if } f_i \neq f_o, G(1) = in, G(2) = out \end{cases}$$

In the first case, interpersonal uncertainty is symmetric for the two individuals, and the ingroup preference factor b alone determines the allocation. $\varepsilon_j \sim N(0, \sigma^2)$ is a normal noise parameter that is i.i.d across observations j. In the second case, the two individuals are symmetric in terms of the group information, and hence, interpersonal uncertainty alone determines the final allocation. In the third case, we have a conjunction of the former two cases, and hence the optimal allocation is a combination of the former two allocations. The parameters a_{IU} and a_{ING} are additive weights that capture the relative importance of interpersonal uncertainty (IU) and ingroup preferences.

The general additive model nests multiple interesting cases. For example, $a_{IU} = a_{ING} = 1$ would imply that the two effects are simply additive. $a_{IU}, a_{ING} > 1$ would imply that the two effects are super-additive, i.e., each effect amplifies in the presence of the other. Similarly, the sub-additivity condition of $a_{IU}, a_{ING} < 1$ would imply that the two effects diminish in the presence of the other. The sub-additive case also includes as relevant special case $a_{IU} + a_{ING} = 1$, where the optimal allocation is a convex combination of both influences. Here, the parameters a_{IU} and a_{ING} collapse into $\alpha = \alpha_{IU} = 1 - \alpha_{ING}$. We label this the *constrained model*.

In our experiment, under the assumption that the valuation of money is measured as WTW, f_1 and f_2 is either the degenerate lottery $(m = 12, 1)$ (tasks) or the 50-50 lottery $(h = 22, .5; l = 4, .5)$. Under these distributions and under CRRA risk-preferences, x_{IU} has a closed form solution. In the case of f_1 being the degenerate and f_2 the 50-50 lottery, the solution is as follows:

$$x_{IU}\left(f_1 = (m, 1), f_2 = (h, 0.5; l, 0.5), \gamma\right) = \frac{M \cdot h\left(\frac{m-l}{h-m}\right)^{\frac{1}{\gamma}} - M \cdot l}{(m-l) + (h-m)\left(\frac{m-l}{h-m}\right)^{\frac{1}{\gamma}}}$$

Using the analytic solutions of x_{IU}, we estimate the parameter vector $\gamma, b, a_{IU}, a_{ING}, \sigma$,

which maximizes the likelihood of the observed data, given our model of x_{1j}.

Results. Table 2 summarizes the results of the structural estimation. Column (1) displays the estimates for the main model, where we estimate a CRRA parameter of $\gamma = 0.374$. To put this in context, we compare the estimate to the mean CRRA parameter estimated in standard risk elicitation tasks. The most closely related elicitation is the Gneezy-Potters (Gneezy and Potters, 1997) investment task, where subjects allocate a fixed sum of $M between a risky and a safe asset with known rates of return. Our setup mimics this task, as our subjects divide a fixed sum between a risky group (WTW is either high or low with equal probability) and a safe group (WTW is medium with certainty). Across 16 studies employing the Gneezy-Potters investment task (1955 subjects in total), Crosetto and Filippin (2016) report an average CRRA parameter of $\gamma = 0.3$. Hence, subjects show on average a slightly higher degree of risk aversion towards interpersonal uncertainty than towards risk in the standard investment task. Next, the structural exercise quantifies the extent of pure ingroup preference, which pulls the optimal allocation away from the 50-50 split and towards the ingroup by $b = 0.741$, or 7% of the total $10. Finally, with respect to the relative importance of interpersonal uncertainty compared to ingroup preferences, we estimate a weight of $a_{IU} = 0.739$ on interpersonal uncertainty and a weight of $a_{ING} = 0.336$ on ingroup preferences. Thus, both factors appear to combine sub-additively in driving ingroup favoritism.

Interestingly, the sum of the individual weights on interpersonal uncertainty and ingroup preferences are close to one ($a_{IU} + a_{ING} = 1.075$). This result motivates our estimation of the constrained model which instead of the two parameters a_{IU} and a_{ING} uses a single parameter α. Compared to the main model, the CRRA and ingroup preference parameters are slightly higher, as displayed in column (2) of Table 2. The relative weight of interpersonal uncertainty is now almost 50%, meaning that both interpersonal uncertainty and ingroup preferences influence allocations to a similar degree. The constrained model provides a worse fit under the Likelihood Ratio test (p-value $< .01$), as also seen by comparing the AIC/ BIC of the two models. This result suggests the main model describes the data more accurately, even accounting for the fact that it contains an additional parameter.

To summarize, our reduced-form and structural analyzes reveal the intensity of people's aversion towards interpersonal uncertainty in the absence of group information and the intensity of group preferences in the absence of interpersonal uncertainty. In the presence of both interpersonal uncertainty and group information, we structurally estimate their respective influence on optimal allocations to be 74% and

34% of their influences in isolation. Thus, the optimal allocation is determined by a sub-additive (close to convex) combination of both influences and not by their sum. This fact helps to interpret our results from the *social* and *non-social* decisions in light of the reduced-form results from the *Ingroup uncertainty* treatment. One might think that the close similarity between the *social* and *non-social* decisions implies that group preferences are irrelevant, which would be at odds not only with the previous literature[22] but also with our results in the *Group information without uncertainty* decision. However, the sub-additivity implies that when one compares a treatment where both factors are present to one where only interpersonal uncertainty matters, the influence of interpersonal uncertainty expands and thus partly compensates for the lack of group preferences. Accordingly, a sub-additive model jointly explains the behavioral patterns of the *social* and *non-social* decisions and the *Ingroup uncertainty* treatment.

4 Self versus other paradigm (Dictator game)

Our experimental design naturally extends to choices involving tradeoffs between one's own utility versus the utility of others (self versus other decisions), as does the idea that interpersonal uncertainty shapes behavior in these tradeoffs.

4.1 Design

Similar to the ingroup versus outgroup case, decision-makers face a *Self social* and a *Self non-social* decision, in randomized order. Right before the *Self non-social* decision, they also complete the valuation task for $100 Amazon gift card money received 6 weeks later.

***Self social* decision.** For the *Self social* decision, we endow decision-makers with $100 which they can allocate between themselves and another individual they have been matched with (without any information about group affiliations). The allocated money is paid out to both parties in the form of Amazon gift card money, six weeks from the date of the experiment. Hence, the *Self social* decision is the standard dictator game: it has consequences for the DM as well as the other individual.

***Self non-social* decision.** In the *Self non-social* decisions, decision-makers split $100 between themselves and another individual, and we remove any social consequences

[22]For instance, findings from the large literature using the minimal-group paradigm (Tajfel and Turner, 1986).

Table 2: Structural estimation results

	Main model (1)	Constrained model (2)
γ	0.374***	0.399***
	(0.016)	(0.013)
b	0.741***	0.911***
	(0.121)	(0.099)
σ	1.881***	1.870***
	(0.046)	(0.046)
a_{IU}	0.739***	
	(0.152)	
a_{ING}	0.336**	
	(0.133)	
α		0.496***
		(0.064)
LL	-1718	-1722
Akaike's IC	3445	3453
Bayesian IC	3469	3472

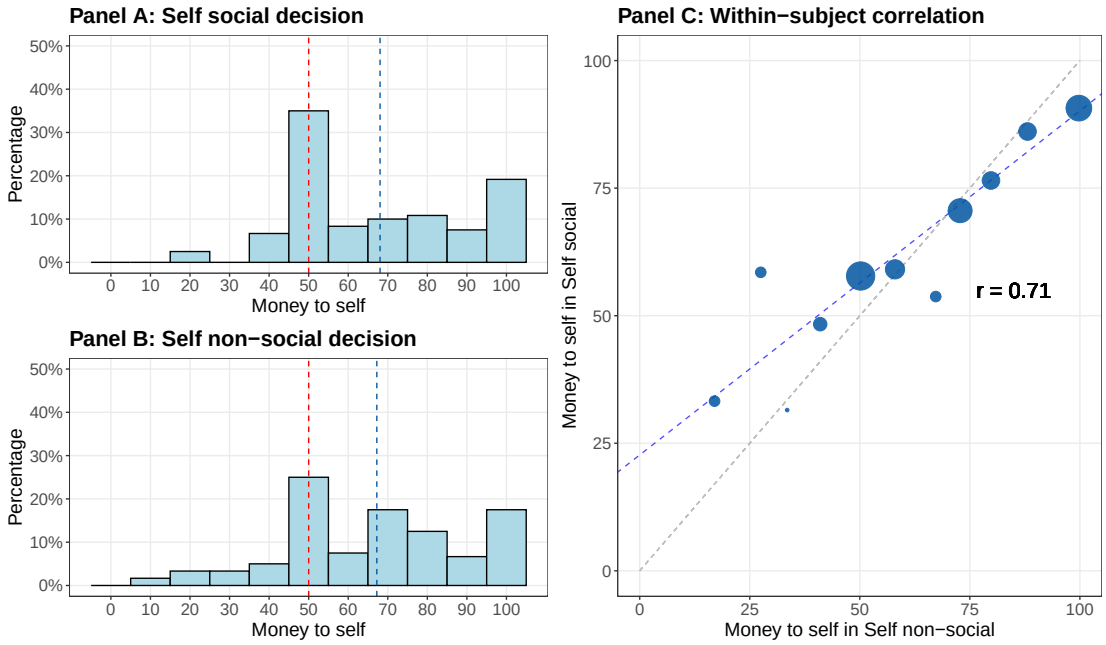
Notes: γ, b, and σ are the CRRA parameter, measure of ingroup preferences, and the standard deviation of the noise term ε respectively. a_{IU} and a_{ING} quantify the importance of interpersonal uncertainty (IU) and ingroup preferences on the optimal allocation choice when both factors are present. The parameter α denotes the relative importance of interpersonal uncertainty relative to ingroup preferences in a model where $\alpha = a_{IU} = 1 - a_{ING}$. See Section 3.5.2 for details. "LL" denotes the maximized Log-Likelihood, "Akaike's IC" is the Akaike's information criterion and "Bayesian IC" the Bayesian information criterion. Significance levels: $^{*}p<0.1$, $^{**}p<0.05$ and $^{***}p<0.01$.

like we did in *ingroup non-social decisions*,. That is, neither the DM, nor the matched participant receive the money that is split. Instead, only the DM receives a reward based on the following formula:

$$\Pi\left(x_{self}, x_{other}\right) = x_{self} \cdot WTP_{self}/100 + x_{other} \cdot WTP_{other}/100$$

where x_{self} and x_{other} are the amounts allocated to self and to the matched individuals respectively, and WTP_{self} and WTP_{other} are their respective WTP for the gift card. Decision-makers are thus incentivized to maximize the sum of their WTP and the WTP of the other individual they are matched with, with both WTPs receiving equal weight. All other elements match the ingroup versus outgroup setting. In total, 120 subjects faced the *Self social* and *Self non-social* decisions.

Figure 3: Main results self versus other decision



Notes: **Panel A and B:** Histogram of the *Self social* (Panel A) and *Self non-social* (Panel B) decision. The x-axis denotes the amount of gift card money (out of \$100) that subjects allocate to themselves. The red dotted line is the even split benchmark, the blue dotted line the average allocation. In *Self social* (Panel A), the decision has consequences for the subjects and the matched other individuals. In *Self non-social* (Panel B), the decision has consequences only for the subjects, with their payoff depending on their and the other individual's WTP for the gift card. **Panel C:** Binned scatter plot of *Self social* and *Self non-social* decisions. The blue dotted line displays the linear fit of a regression of the *Self social* on the *Self non-social* decision. The correlation coefficient is $r = 0.71$. For all three panels, the binwidth is 10. Displayed are $n = 61$ decisions by 61 subjects in Panel A, $n = 59$ decisions by 59 subjects in Panel B, and $n = 120$ decision-pairs by 120 subjects in Panel C.

4.2 Results

***Self social* decision.** In the *Self social* decision, subjects allocate on average \$69.05 to themselves, thereby allocating significantly more money to themselves compared to the equal split ($p < 0.001$, one-sample Wilcoxon test). Figure 3 panel A displays the distribution, which replicates the typical distributional pattern of dictator games found in the literature (e.g., Engel, 2011). In total, 61% of subjects allocate more money to themselves, 3% allocate more to the other person, and 36% implement the 50/50 split.

***Self non-social* decision.** In *Self non-social*, subjects allocate on average \$64.12 to themselves, again a significant deviation from the equal split ($p < 0.001$, one-sample Wilcoxon test). As Figure 3 panel B shows, the distribution is also similarly shaped as in the *Self social* case. In total, 58% of subjects allocate more money to themselves, 17% allocate more to the other person, and 25% implement the 50/50 split.

Comparing *Self social* and *non-social*. Allocations in the *Self non-social* setting closely replicate the behavior we observe in *Self social*. Statistically, we cannot reject that the average amount that subjects allocate to themselves is equal across *social* and *non-social* decisions ($p = 0.27$, unpaired Wilcoxon test). Similarly, we cannot reject that the distribution of allocations is equal across the two decisions ($p = 0.39$, Kolmogorov-Smirnov test).

These results also hold in the within-subject comparison. Figure 3 panel C binscatter-plots each individual's social and non-social decision pair. The two decisions are highly correlated at the individual level, with a correlation coefficient of $r = 0.71$. Hence, the *Self non-social* decision strongly predicts the *Self social* decision.

Robustness. To show that people are attentive to the *non-social* incentives in the self versus other setting, we use the same the incentive treatment as in the ingroup versus outgroup setting (Section 3.3) and apply it to the current setting. Our results are similar: subjects understand the incentives and react to them as hypothesized. We provide the results in Appendix Section F.

Result 2. *Self non-social choices replicate the self-favoring behavior found in Self social choices. The distributions are similar and strongly correlated on the individual level.*

4.3 Relating interpersonal uncertainty to dictator game giving

After the *Self social* decision, subjects reported self-reported interpersonal uncertainty on a 0 to 10-point Likert-scale both over their own valuation and their perception of the other individual's valuation. Subjects report on average 2.59 Likert-scale points higher uncertainty for the other person's valuation than their own, which is again significantly different from the no difference benchmark ($p < 0.001$, paired Wilcoxon tests).[23] In total, 72% of subjects report a higher uncertainty about the other person's valuations, 6% report a higher uncertainty about their own valuation, and for 23% report equal ratings. Importantly, the difference in ratings again predicts choice behavior. Higher uncertainty ratings of the other individual are associated with subjects allocating more money to themselves, e.g., displaying less other-regarding behavior in the *Self social* $(r = 0.35, p < 0.001)$ and *Self non-social* decision $(r = 0.24, p < 0.01)$. See Appendix Figure B.5 for a binned scatter plot.

[23] As expected, subjects report low degrees of uncertainty over their own valuation for the future gift card money, with a median report of 1 on the 0 to 10 Likert-scale. See Chakraborty (2021) and Gabaix and Laibson (2022) for models where DMs are uncertain about their own future utility.

Robustness. To ensure that our uncertainty measure was not confounded by prior *social* or *non-social* decisions, we conducted a robustness treatment (*Self/other belief measurement*, $n = 61$) where subjects only reported their uncertainty about own and other's valuation on the 11-point Likert-scale. We replicate the previously reported results: subjects perceived 3.11 Likert-scale points higher uncertainty over the other person's valuation compared uncertainty about their own ($p < 0.001$, paired Wilcoxon test). In total, 84% of subjects report higher uncertainty about others' valuations, 10% report the same degree of uncertainty, and 7% report higher uncertainty about their own valuation.

We also asked them to estimate the mean WTP of other subjects. On average, subjects report a mean WTP of \$83.43 and on average have a WTP of \$86.69. Thus, on average, subjects believe others to have a lower WTP than themselves ($p = 0.03$, paired Wilcoxon test). Overall, for 52% of subjects, their own WTP is higher than their estimate of others' average, the reverse is true for 41%, and for 7% they coincide. Accordingly, if the modal subject was a pure utilitarian who altruistically treats others equally as himself, he would still display self-favoritism. This result echoes our central argument that not accounting for interpersonal uncertainty in self-favoritism might result in underestimating the extent of altruism.

5 Giving versus taking paradigm

Next, we turn to studying how mean-shifted interpersonal uncertainty influences redistribution behavior. The previous literature primarily finds that redistribution behavior is merit-based: people redistribute less from initial endowments if these endowments are earned compared to generated by chance (Cappelen, Falch, and Tungodden, 2020). In particular, in the context of dictator games, several studies show that if the initial endowment was earned instead of being windfall, then dictators increase their allocation towards the individual earning the endowment. (Ruffle, 1998; Cherry, 2001; Cherry, Frykblom, and Shogren, 2002; Cherry and Shogren, 2008; Oxoby and Spraggon, 2008; Krupka and Weber, 2013).

This behavior is typically attributed to fairness preferences (e.g., Tungodden and Cappelen, 2019), fairness-based social norms (Krupka and Weber, 2013), or the role of property rights (Oxoby and Spraggon, 2008). We offer an alternative explanation: if people perceive that on average, the disutility from losing earned money exceeds the utility from gaining money (i.e, a gain-loss asymmetry), then dictators would perceive mean-shifted interpersonal uncertainty for recipients who have earned the endowment compared to recipients who have not. Then, based on our framework

and our result (iii) of Theorem 1, a simple utilitarian motive under uncertainty would also lead to the same asymmetry between giving and taking environments. Our next treatments test this channel of mean-shifted interpersonal uncertainty.

5.1 Design

Following the typical setup of the literature, we alter our dictator game from a giving environment to a taking environment. DMs face a *Taking social* decision and a *Taking non-social* decision. In both *Taking* decisions, DMs are matched to a previous participant who has earned $100 for participating in a previous study, scheduled to be paid in 6 weeks from the study day. In total, 123 subjects participated in this experiment.

***Taking social* decision.** In the *social* variant, the DM decides whether to take some or all of the money that the other participant has earned for themselves, adapting the design of Oxoby and Spraggon (2008). The chosen allocation is then implemented with consequences for the DM and the other participant.

***Taking non-social* decision.** In the *non-social* variant, we replicate the setup described in section 3 with one key difference: because the other participant already earned the $100 that was up for splitting, the DM's utilitarian incentives were calculated using the other participant's willingness-to-accept (WTA) for gift card money, instead of their WTP. Thus the DM's payment depended on their own WTP and the matched participant's WTA. Specifically, the incentive for the DM is as follows:

$$\Pi\left(x_{self}, x_{other}\right) = x_{self} \cdot WTP_{self}/100 + x_{other} \cdot WTA_{other}/100$$

with x_{self} and x_{other} denoting the money DMs allocate to themselves and the other individual respectively, WTP_{self} is their own WTP and WTA_{other} is the other individual's WTA for the gift card money.

After the DMs participated in the MPL that elicits their WTP, we explained to them the following details about matched participants: First, the matched participants earned the $100 gift card that would pay in 6 weeks, through their participation. Then, we asked them whether they would be willing to give away the gift card in exchange for an immediately payable monetary amount. We ask this question for different amounts of the immediately payable money, using an MPL, to elicit their WTA. The DMs are already familiar with the MPL-elicitation method at this point. We emphasize to DMs that the only difference between their's and the matched participant's elicitation is, instead of having the option to receive the gift card, the matched participants already 'owned' the gift card and had the opportunity to sell it.

Multiple studies have found that WTA is, on average, higher than WTP (see Camerer, 1995; DellaVigna, 2009, for overviews), and hence WTA>WTP is a well-established empirical pattern. Our central hypothesis is that, if DMs also anticipate the WTA-WTP gap as mean-shifted interpersonal uncertainty, then utilitarianism provides a novel foundation for differences in giving and taking (Theorem 1). In particular, under WTA>WTP, we predict that the amount allocated to the matched participant should increase (compared to the giving paradigm) not only in *Taking social*, but also in *Taking non-social*. Further, because *Taking non-social* does not feature any scope for fairness attitudes, we can separate our channel from a fairness channel.

Table 3: Dictator game allocations under giving setting versus taking setting

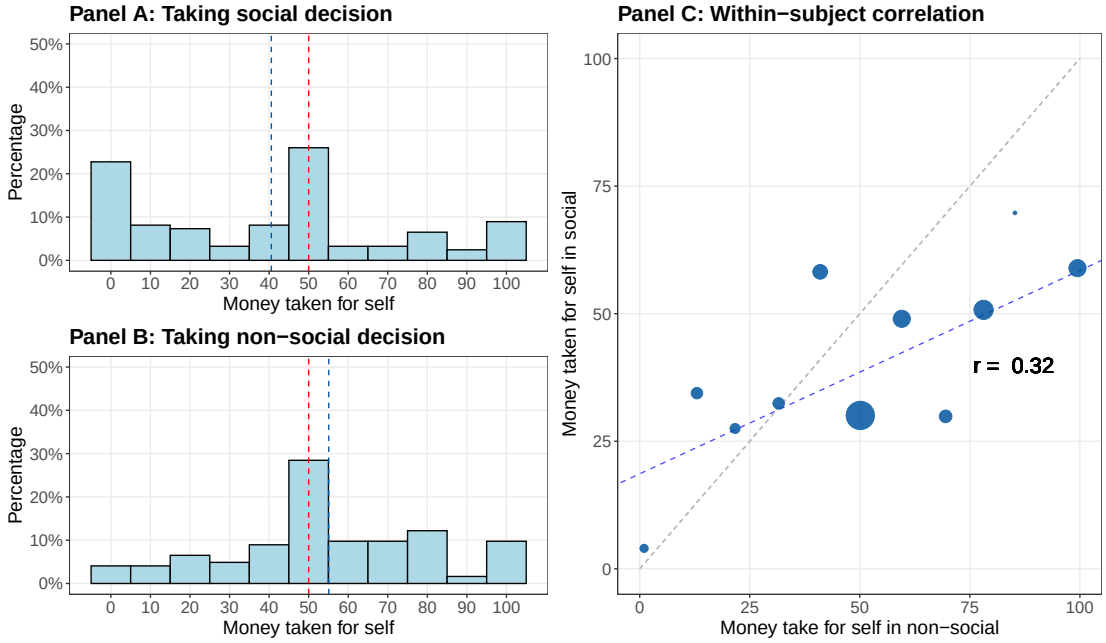
	Dependent variable:	
	Allocation to self	
	Social decision	*Non-social decision*
	(1)	(2)
Constant (Giving setting)	68.050***	66.193***
	(2.719)	(2.634)
Taking setting	−27.481***	−12.077***
	(3.529)	(3.114)
Order: *social* decision first	−0.001	2.046
	(3.545)	(3.114)
Observations	243	243
R^2	0.201	0.061

Notes: The table shows OLS estimates. The dependent variables in (1) and (2) are the amount subjects allocate to themselves (out of $100) in the *Self social* decisions and in the *Self non-social* decisions respectively. "Taking setting" is an indicator variable equal to one if the allocation decision concerns taking earned money away from the matched participant. "Order: *social* decision first" is an indicator variable equal to one if subjects faced the *social* decision before facing the *non-social* decision. Robust standard errors in parentheses. Significance levels: *p<0.1, **p<0.05 and ***p<0.01.

5.2 Results

Comparing *Taking* to the *Giving* setting. Table 3 displays our main result: subjects allocate significantly less money to themselves in the *Taking* compared to the *Giving* setting of Section 4. Comparing *Self social* with *Taking social* decisions, displayed in column (1), we see a significant decrease of $27.48 in the amount subjects

Figure 4: Giving versus taking results



Notes: **Panel A and B:** Histogram of the *Taking social* (Panel A) and *Taking non-social* (Panel B) decision. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to themselves instead of another individual. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Taking social* (Panel A), the other individual has earned the $100 and subjects decide how much to take away for themselves. Their decision has consequences for themselves and the other individual. In *Taking non-social* (Panel B), the decision has consequences only for the subjects, with their payoff depending on their WTP and the other individual's WTA for the gift card. **Panel C:** Binned scatter plot of the *Taking social* and *Taking non-social* decision. The blue dotted line displays the linear fit of a regression of the *Taking social* on the *Taking non-social* decision. The correlation coefficient is $r = 0.32$. For all three panels, the binwidth is 10. Displayed are $n = 123$ decision-pairs by 123 subjects.

allocate to themselves, using the within-subject data controlling for the order. We thus replicate the common finding of aversion to taking from earned endowments in the literature with our *social* decisions. In column (2), we compare *Self non-social* with *Taking non-social* decisions, where we also find a significant decrease of $12.08 in the amount subjects allocate to themselves. Therefore, incentivizing DMs with the other individuals' WTA instead of their WTP induces DMs to allocate less to themselves. The decrease in allocation to the self from the *Giving* to the *Taking* setting is markedly smaller in the *non-social* case compared to the *social* case.

Comparing *Taking social* and *Taking non-social*. Figure 4 displays the comparison of the *Taking social* and *Taking non-social* decisions. Panel A shows the distribution of choices in the *Taking social* decision, where 26% of subjects allocate more money to themselves, 51% allocate more to the other person, with the remaining 23% allocating the even split. In the *Taking non-social* decision, displayed in Panel

B, 46% allocate more money to themselves, 29% allocate more to the other person, and 24% split evenly. We see a significant within-subject correlation of $r = 0.32$ ($p < 0.001$) between *Taking social* and *Taking non-social*, see Panel C of Figure 4. Thus, taking behavior in the *social* decision correlates with the *non-social* decision that does not feature taking (not even in how the instructions were framed). Still, contrary to the other settings, these decisions differ in average allocations ($p < 0.001$, unpaired Wilcoxon test), and distributions ($p < 0.01$, Kolmogorov-Smirnov test). These results suggest that *social* decisions are also driven by motives that are absent in the *non-social* decisions.

A potential motive comes from the observation that 22% of subjects choose to take \$0 for themselves in *Taking social*, while only 3% do so in *Taking non-social*. In contrast, in *Self social* and *Self non-social*, not a single subject chooses to give everything to the other individual. This pattern suggests that fairness preferences are also at work, e.g., some subjects have a strong libertarian fairness view (Alma, Cappelen, and Tungodden, 2020) or adhere to a deontological motive that, independent of consequences, it is not permissible to take money from someone (Bénabou, Falk, and Henkel, 2024). Interestingly, those subjects refusing to take any money completely explain the gap between *Taking social* and *Taking non-social*. If we focus only on subjects who take more than \$0 for themselves in *Taking social*, we can no longer reject the equality of average giving between *Taking social* and *Taking non-social* ($p = 0.22$, unpaired Wilcoxon test)[24] and distributions become more similar ($p = 0.09$, Kolmogorov-Smirnov test), see Appendix Figure B.6. Similarly, the within-subject correlation increases to $r = 0.45$.

A necessary condition for interpreting our hypotheses and results is that subjects indeed perceive a positive difference in the utility impact of taking earned money and giving windfall money, i.e., in other's WTA and WTP. To validate this assumption, we asked subjects whether they generally think that a person's WTA for the gift card is higher, lower or equal to the WTP. In total, 46% of subjects believe WTA to be higher than WTP, 29% believe WTP to be higher, and 24% believe both to be equal. Thus, subjects believe WTA>WTP on average.

Robustness. For robustness, subjects in the *Self/other belief measurement* treatment ($n = 61$), who reported the mean and the uncertainty over other's WTP valuations, also reported the same quantities about other's WTA valuations. Thus, each subject reported their estimated mean and their perceived uncertainty over other's WTP/

[24]Note that this effect is not mechanical because we remove both *social* and *non-social* decisions due to the within-subject structure of our data.

WTA . On average, subjects reported the mean WTA to be $8.00 higher than the mean WTP ($91.43 compared to $83.43), a significant difference ($p < 0.001$, paired Wilcoxon-test). In total, 64% of subjects estimated the mean WTA to be higher than the mean WTP, 20% the estimated the reverse, and 16% estimated both to be equal.

With respect to uncertainty, subjects report a 0.51 Likert scale points higher uncertainty about others' WTP than about others' WTA, a significant difference ($p = 0.02$, paired Wilcoxon-test). These results suggests that both mean-shifted interpersonal uncertainty and differences in perceived interpersonal uncertainty contribute to the giving versus taking differences.

Result 3. *Subjects allocate more money to the other person when allocating the other person's earned money (Taking social) than when allocating windfall money (Self social). The allocations are ranked similarly in Self non-social with Taking non-social.*

6 Conclusion

In this paper, we provide a conceptual framework that explains how people's response to uncertainty influences prominent features of prosocial behavior. We design and implement a series of experiments documenting how uncertainty can bolster ingroup favoritism, weaken altruistic giving, and shape redistributive behavior based on the source of endowment. Our results thus challenge the notion that prosocial decisions are purely a measure of social preferences. Instead, our results suggest that a significant degree of heterogeneity in prosocial behavior, both within a given decision setting and between different settings, is driven by people's differential response to interpersonal uncertainty.

As a consequence, precise identification of social preferences from prosocial behavior requires explicit accounting for interpersonal uncertainty. Otherwise, depending on the nature of interpersonal uncertainty, parameters of social preferences may be over- or underestimated. We also demonstrate an experimental design to disentangle uncertainty from preferences: a researcher can exogenously vary interpersonal uncertainty to explicitly measure and control for it. For instance, in our experiment, we provide subjects with information so that interpersonal uncertainty switches from favoring one to the other recipient, or is balanced among recipients.

Finally, our framework suggests that prosocial behavior is malleable and interpersonal uncertainty can explain the dynamics of prosocial behavior. Indeed, a large literature has shown evidence that patterns such as intergroup behavior and attitudes are not necessarily fixed across time and individuals but change in response

to interventions, context, and experiences. For instance, intergroup contact created by spatial proximity (Bursztyn et al., 2024), shared classes (Rao, 2019), shared living (Corno, La Ferrara, and Burns, 2022), sports events (Mousa, 2020; Lowe, 2021) and attending youth camps (Ghosh et al., 2024) reduces intergroup frictions and ingroup favoritism. Our framework suggests that contact reduces the interpersonal uncertainty about the outgroup, both in absolute and relative terms. As a result, contact can change behavior towards the outgroup even without changing group preferences. Similarly, our conceptual framework and results vindicate how people's degree of favoritism towards specific groups varies based on their closeness (Fong and Luttmer, 2009), salience of shared experiences (McLeish and Oxoby, 2011), or (perceived) similarity (Goeree et al., 2010; Fouka, Mazumder, and Tabellini, 2022) to ingroup members. As such, targeting and reducing interpersonal uncertainty could be a promising avenue for fostering more prosocial behavior, bridging animosities, and decreasing intergroup conflict.

References

Alma, Ingvild, Alexander W Cappelen, and Bertil Tungodden (2020). "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5): 1753–88. [40]

Amador, Manuel, Ivan Werning, and George-Marios Angeletos (2006). "Commitment vs. Flexibility." *Econometrica* 74 (2): 365–96. [7]

Augenblick, Ned, Muriel Niederle, and Charles Sprenger (2015). "Working Over Time: Dynamic Inconsistency in Real Effort Tasks." *Quarterly Journal of Economics* 130 (3): 1067–115. [24]

Bénabou, Roland, Armin Falk, and Luca Henkel (2024). "Ends versus Means: Kantians, Utilitarians and Moral Decisions." *Working Paper*, [40]

Berg, Joyce, John W. Dickhaut, and Kevin McCabe (1995). "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–42. [7]

Böhm, Robert, Hannes Rusch, and Jonathan Baron (2020). "The Psychology of Intergroup Conflict: A Review of Theories and Measures." *Journal of Economic Behavior and Organization* 178: 947–62. [6]

Brock, J. Michelle, Andreas Lange, and Erkut Y Ozbay (2013). "Dictating the Risk: Experimental Evidence on Giving in Risky Environments." *American Economic Review* 103 (1): 415–37. [7]

Bursztyn, Leonardo, Thomas Chaney, Tarek A. Hassan, and Aakaash Rao (2024). "The Immigrant Next Door." *American Economic Review* 114 (2): 348–84. [42]

Camerer, Colin F. (1995). "Individual Decision Making." In *Handbook of Experimental Economics*. John H. Kagel and Alvin E. Roth, ed. vol. 2, Princeton, NJ: Princeton University Press, 587–703. [38]

Cappelen, Alexander W, Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden (2022). "A Meritocratic Origin of Egalitarian Behaviour." *Economic Journal* 132 (646): 2101–17. [7]

Cappelen, Alexander W., Thomas De Haan, and Bertil Tungodden (2024). "Fairness and Limited Information: Are People Bayesian Meritocrats?" *Journal of Public Economics* 233 (5): 105097. [7]

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden (2020). "Fair and Unfair Income Inequality." In *Handbook of Labor, Human Resources and Population Economics*. Klaus F. Zimmermann, ed. Cham: Springer International Publishing, 1–25. [6, 36]

Capraro, Valerio, Joseph Y. Halpern, and Matjaz Perc (2024). "From Outcome-Based to Language-Based Preferences." *Journal of Economic Literature* 62 (1): 115–54. [6, 11]

Cettolin, Elena, Arno Riedl, and Giang Tran (2017). "Giving in the Face of Risk." *Journal of Risk and Uncertainty* 55 (2-3): 95–118. [7]

Chakraborty, Anujit (2021). "Present bias." *Econometrica* 89 (4): 1921–61. [7, 35]

Chakraborty, Anujit, Yoram Halevy, and Kota Saito (2020). "The relation between behavior under risk and over time." *American Economic Review: Insights* 2 (1): 1–16. [7]

Charness, Gary, and Yan Chen (2020). "Social Identity, Group Behavior, and Teams." *Annual Review of Economics* 12: 691–713. [6]

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree-An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97. [14]

Chen, Yan, and Sherry Xin Li (2009). "Group Identity and Social Preferences." *American Economic Review* 99 (1): 431–57. [11]

Cherry, Todd L, Peter Frykblom, and Jason F Shogren (2002). "Hardnose the Dictator." *American Economic Review* 92 (4): 1218–21. [7, 36]

Cherry, Todd L. (2001). "Mental Accounting and Other-Regarding Behavior: Evidence from the Lab." *Journal of Economic Psychology* 22 (5): 605–15. [7, 36]

Cherry, Todd L., and Jason F. Shogren (2008). "Self-Interest, Sympathy and the Origin of Endowments." *Economics Letters* 101 (1): 69–72. [7, 36]

Corno, Lucia, Eliana La Ferrara, and Justine Burns (2022). "Interaction, Stereotypes, and Performance: Evidence from South Africa." *American Economic Review* 112 (12): 3848–75. [42]

Crosetto, Paolo, and Antonio Filippin (2016). "A theoretical and experimental appraisal of four risk elicitation methods." *Experimental Economics* 19: 613–41. [31]

Dasgupta, Partha, and Eric Maskin (2005). "Uncertainty and hyperbolic discounting." *American Economic Review* 95 (4): 1290–99. [7]

DellaVigna, Stefano (2009). "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature* 47 (2): 315–72. [38]

Engel, Christoph (2011). "Dictator Games: A Meta Study." *Experimental Economics* 14 (4): 583–610. [34]

Enke, Benjamin (2023). "Moral Boundaries." *Annual Review of Economics* forthcoming: [6]

Enke, Benjamin, and Thomas Graeber (2023). "Cognitive Uncertainty." *Quarterly Journal of Economics* 138 (4): 2021–67. [7]

Enke, Benjamin, Thomas Graeber, and Ryan Oprea (2023). "Complexity and Time." *Working Paper*, [7]

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann (2022). "Moral Universalism: Measurement and Economic Relevance." *Management Science* 68 (5): 3590–603. [11, 12, 15]

Exley, Christine L. (2016). "Excusing Selfishness in Charitable Giving: The Role of Risk." *Review of Economic Studies* 83 (2): 587–628. [7]

Eyal, Peer, Rothschild David, Gordon Andrew, Evernden Zak, and Damer Ekaterina (2021). "Data Quality of Platforms and Panels for Online Behavioral Research." *Behavior Research Methods* 54: 1643–62. [14]

Falk, Armin, Ernst Fehr, and Urs Fischbacher (2008). "Testing Theories of Fairness-Intentions Matter." *Games and Economic Behavior* 62 (1): 287–303. [7]

Fehr, Ernst, and Gary Charness (2023). "Social Preferences: Fundamental Characteristics and Economic Consequences." *Working Paper*, [6]

Fehr, Ernst, and Simon Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14 (3): 159–81. [7]

Fong, Christina M, and Erzo F. P Luttmer (2009). "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty." *American Economic Journal: Applied Economics* 1 (2): 64–87. [42]

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini (2022). "From Immigrants to Americans: Race and Assimilation during the Great Migration." *Review of Economic Studies* 89 (2): 811–42. [42]

Gabaix, Xavier, and David Laibson (2022). "Myopia and Discounting." *Working Paper*, [35]

Ghosh, Arkadev, Prerna Kundu, Matt Lowe, and Gareth Nellis (2024). "Creating Cohesive Communities: A Youth Camp Experiment in India." *Working Paper*, [42]

Gill, David, and Victoria Prowse (2012). "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review* 102 (1): 469–503. [20]

Gill, David, and Victoria Prowse (2019). "Measuring Costly Effort Using the Slider Task." *Journal of Behavioral and Experimental Finance* 21 (3): 1–9. [20, 24]

Gneezy, Uri, and Jan Potters (1997). "An experiment on risk taking and evaluation periods." *quarterly journal of economics* 112 (2): 631–45. [31]

Goeree, Jacob K, Margaret A McConnell, Tiffany Mitchell, Tracey Tromp, and Leeat Yariv (2010). "The 1/ *d* Law of Giving." *American Economic Journal: Microeconomics* 2 (1): 183–203. [42]

Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto (2013). "Moral Foundations Theory." In *Advances in Experimental Social Psychology*. vol. 47, Elsevier, 55–130. [11]

Gupta, Neeraja, Luca Rigotti, and Alistair Wilson (2021). "The Experimenters' Dilemma: Inferential Preferences over Populations." *Working Paper*, [14]

Halevy, Yoram (2008). "Strotz meets Allais: Diminishing impatience and the certainty effect." *American Economic Review* 98 (3): 1145–62. [7]

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood (2019). "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22 (1): 129–46. [6]

Krupka, Erin L., and Roberto A. Weber (2013). "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?" *Journal of the European Economic Association* 11 (3): 495–524. [7, 36]

Lowe, Matt (2021). "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *American Economic Review* 111 (6): 1807–44. [42]

McLeish, Kendra N., and Robert J. Oxoby (2011). "Social Interactions and the Salience of Social Identity." *Journal of Economic Psychology* 32 (1): 172–78. [42]

Mousa, Salma (2020). "Building Social Cohesion between Christians and Muslims through Soccer in Post-ISIS Iraq." *Science* 369 (6505): 866–70. [42]

Oprea, Ryan (2024). "Simplicity Equivalents." *Working Paper*, [7]

Oxoby, Robert J., and John Spraggon (2008). "Mine and Yours: Property Rights in Dictator Games." *Journal of Economic Behavior & Organization* 65 (3-4): 703–13. [7, 36, 37]

Rao, Gautam (2019). "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools." *American Economic Review* 109 (3): 774–809. [42]

Ruffle, Bradley J. (1998). "More Is Better, But Fair Is Fair: Tipping in Dictator and Ultimatum Games." *Games and Economic Behavior* 23 (2): 247–65. [7, 36]

Shayo, Moses (2020). "Social Identity and Economic Policy." *Annual Review of Economics* 12: 355–89. [6]

Sozou, Peter D (1998). "On hyperbolic discounting and uncertain hazard rates." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265 (1409): 2015–20. [7]

Tabellini, Guido (2008). "The Scope of Cooperation: Values and Incentives." *Quarterly Journal of Economics* 123 (3): 905–50. [11]

Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament (1971). "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1 (2): 149–78. [11]

Tajfel, Henri, and John Turner (1986). "The Social Identity Theory of Intergroup Behavior." In *The Psychology of Intergroup Relations*. Stephen Worchel and William Austin, ed. Chicago: Nelson-Hall, 7–24. [32]

Tungodden, Bertil, and Alexander W. Cappelen, eds. (2019). *The Economics of Fairness*. Edward Elgar Publishing. [36]

Appendix

A Proofs

Proof of Theorem 1. For simplicity, whenever possible we will write $E_{v_i \sim f_i}$ simply as E_{f_i}. From the utility expression, we get

$$\frac{d}{dx} EU(x) = E_{f_1, f_2} \left((v_1 - v_2) U' \left(v_1 x + v_2 (100 - x) \right) \right)$$

and

$$\frac{d^2}{dx^2} EU(x) = E_{f_1, f_2} \left((v_1 - v_2)^2 U'' \left(v_1 x + v_2 (100 - x) \right) \right) < 0$$

as $U'' < 0$ and $f_1, f_2 \geq 0$. $\frac{d^2}{dx^2} EU(x)$ being strictly positive implies that $\frac{d}{dx} EU(x) = 0$ must be obtained at a unique point. Evaluating the first derivative at $x = 50$, we get

$$\frac{d}{dx} EU(x)|_{x=50} = E_{f_1, f_2} (v_1 - v_2) U' \left(50 v_1 + 50 v_2 \right) \tag{2}$$

(i) When f_1 and f_2 are identical, then we can also rewrite

$$\begin{aligned} \frac{d}{dx} EU(x) &= E_{v_1 \sim f_2} E_{v_2 \sim f_1} (v_1 - v_2) U' \left(v_1 x + v_2 (100 - x) \right) \\ &= E_{v_2 \sim f_2} E_{v_1 \sim f_1} (v_2 - v_1) U' \left(v_2 x + v_1 (100 - x) \right) \\ &= E_{f_1, f_2} (v_2 - v_1) U' \left(v_2 x + v_1 (100 - x) \right) \end{aligned}$$

where the first step integrates v_1 over f_2 and v_2 over f_1 instead, the second step interchanges the names of variables (v_1 and v_2) of integration, and the last step interchanges the order of integration. Now evaluating the final expression at $x = 50$, we get

$$\frac{d}{dx} EU(x)|_{x=50} = E_{f_1, f_2} (v_2 - v_1) U' \left(50 v_1 + 50 v_2 \right) \tag{3}$$

Equations 2 and 3 together imply $\frac{d}{dx} EU(x)|_{x=50} = -\frac{d}{dx} EU(x)|_{x=50} = 0$.

ii) When f_2 is a mean-preserving spread of f_1, then there exists a random variable $z \sim f_z$ with zero expectation conditional on any given value of v_1, such that v_2 has the same distribution as $v_1 + z$, or in other words, $v_2 =^d v_1 + z$. Therefore, we can replace v_2 by a variable $w_1 + z$ where w_1 and v_1 both have identical distribution f_1.

$$\frac{d}{dx} EU(x) = E_{v_1 \sim f_1} E_{w_1 \sim f_1} E_{z | v_1, w_1} (v_1 - w_1 - z) U' \left(v_1 x + (w_1 + z)(100 - x) \right)$$

Therefore,

$$\frac{d}{dx}EU(x)|_{x=50} = E_{f_1}E_{f_1}E_{z|v_1,w_1}(v_1 - w_1 - z)U'\left(50v_1 + 50w_1 + 50z\right) \tag{4}$$

Because $E_{f_1}E_{f_1}$ is integrating with respect to two identical independent distributions, we can interchange their variable names (w_1 and v_1) in Equation 4:

$$\frac{d}{dx}EU(x)|_{x=50} = E_{f_1,f_1}E_{z|v_1,w_1}(w_1 - v_1 - z)U'\left(50v_1 + 50w_1 + 50z\right) \tag{5}$$

Adding equations 4 and 5, and then using law of iterated expectations:

$$\begin{aligned}
2\frac{d}{dx}EU(x)|_{x=50} &= -E_{f_1,f_1}\left(E_{z|v_1,w_1}2zU'\left(50v_1 + 50w_1 + 50z\right)\right) \\
&> -E_{f_1,f_1}\left(E_{z|v_1,w_1}2zU'\left(50v_1 + 50w_1\right)\right) \\
&= -E_{f_1,f_1}U'\left(50v_1 + 50w_1\right)\left(E_{z|v_1,w_1}2zf_z(z|v_1)dz\right) \\
&= 0
\end{aligned}$$

The inequality uses the fact: $zU'\left(50v_1 + 50w_1 + 50z\right) < zU'\left(50v_1 + 50w_1\right)$ for both $z > 0$ and $z < 0$. The last step follows from the fact that $E_{z|v_1,w_1}z = 0$. Therefore, $\frac{d}{dx}EU(x)|_{x=50} > 0$, and thus, the optimal allocation $x^* > 50$. Next,

$$\begin{aligned}
\frac{d}{dx}EU(x)|_{x=100} &= E_{f_1}E_{f_1}E_{z|v_1,w_1}(v_1 - w_1 - z)U'\left(100v_1\right) \\
&= E_{f_1}E_{f_1}U'\left(100v_1\right)E_{z|v_1,w_1}(v_1 - w_1 - z) \\
&= E_{f_1}E_{f_1}U'\left(100v_1\right)(v_1 - w_1) \\
&= E_{f_1}E_{f_1}U'\left(100w_1\right)(w_1 - v_1) \\
&= \frac{1}{2}E_{f_1}E_{f_1}[U'\left(100w_1\right)(w_1 - v_1) + U'\left(100v_1\right)(v_1 - w_1)] \\
&= \frac{1}{2}E_{f_1}E_{f_1}(U'(100w_1) - U'(100v_1))(w_1 - v_1) \\
&< 0
\end{aligned}$$

The first step replaces $x = 100$ into the expression of $\frac{d}{dx}EU(x)$ derived at the beginning of the proof. The second step uses that $U'\left(100v_1\right)$ is independent of z. The third step uses $E_{z|v_1,w_1}z = 0$. The fourth step uses the property that v_1, w_1 are drawn i.i.d from f_1, and hence those two variable names can be interchanged. The fifth step uses the average of the two expressions from the previous lines. The last step uses the property that U' is decreasing.

As $\frac{d}{dx}EU(x)|_{x=100} < 0$, the concavity of the expression implies that $\frac{d}{dx}EU(x) = 0$ must be obtained at some $50 < x < 100$.

(iii)

The first derivative of the objective function, evaluated at x^*, should be zero.

$$E_{f_1,f_2}(v_1 - v_2 - c)U'(x^*v_1 + (100 - x^*)(v_2 + c)) = 0 \tag{6}$$

First, taking the implicit derivative of the last equation w.r.t c, and then re-arranging:

$$E_{f_1,f_2}[-U' + (v_1 - v_2 - c)^2\frac{dx^*}{dc}U'' + (v_1 - v_2 - c)(100 - x^*)U''] = 0$$

Next, we re-arrange and then bound $\frac{dx^*}{dc}$ in 6 steps as explained below. Under CARA,

$$\begin{aligned}
\frac{dx^*}{dc} &= \frac{E_{f_1,f_2} - U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} + \frac{E_{f_1,f_2}(v_1 - v_2 - c)(100 - x^*)U''}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2} - U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} + \frac{E_{f_1,f_2}(v_1 - v_2 - c)(100 - x^*) \times \frac{U''_{100}}{U'_{100}}U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2} - U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} + \frac{(U''_{100})(100 - x^*)}{U'_{100}}\frac{E_{f_1,f_2}(v_1 - v_2 - c)U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2} - U'}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} + \frac{(U''_{100})(100 - x^*)}{U'_{100}} \times \frac{0}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&< 0
\end{aligned}$$

The second step utilizes the assumption of constant absolute risk aversion: $\frac{U''}{U'} = \frac{U''_{100}}{U'_{100}}$, and hence, $U'' = \frac{U''_{100}}{U'_{100}}U'$. The third step simply reorganizes the numerator in the second additive term. The fourth step uses equation 6 to set $E_{f_1,f_2}(v_1 - v_2 - c)U'$ to zero. The last step uses $U' > 0, U'' < 0$.

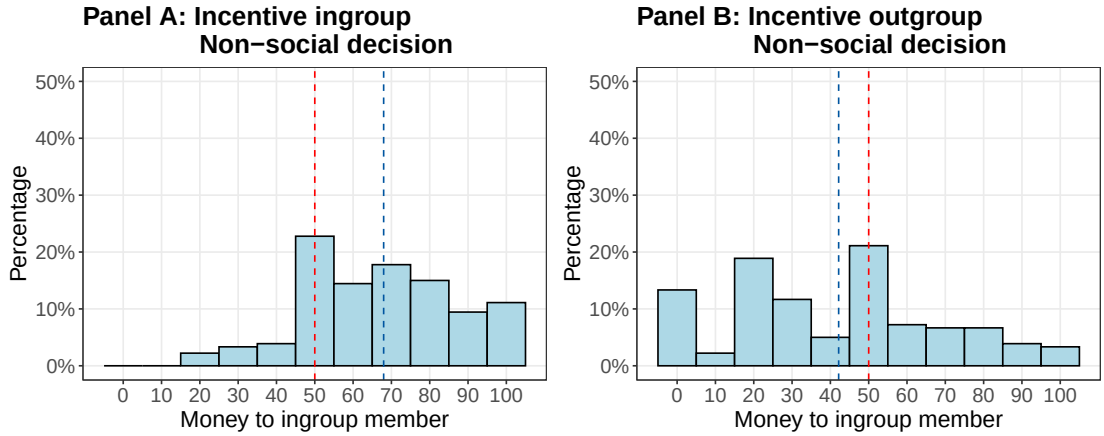
Under CRRA preferences,

$$\begin{aligned}
\frac{dx^*}{dc} &= \frac{E_{f_1,f_2}[-U' + (v_1 - v_2 - c)(100 - x^*)U'']}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2}[-U' - (x^*v_1 + (100 - x)(v_2 + c))U'' + 100v_1U'']}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2}[-U' + rU' + 100v_1U'']}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''} \\
&= \frac{E_{f_1,f_2}[-(1 - r)U' + 100v_1U'']}{-E_{f_1,f_2}(v_1 + c - v_2)^2U''}
\end{aligned}$$

The third step utilizes the CRRA parameter $r < 1$. In the last expression, the numerator is negative as $r < 1, v_1 \geq 0, U'' < 0$ and the denominator is positive, which concluces the proof. □

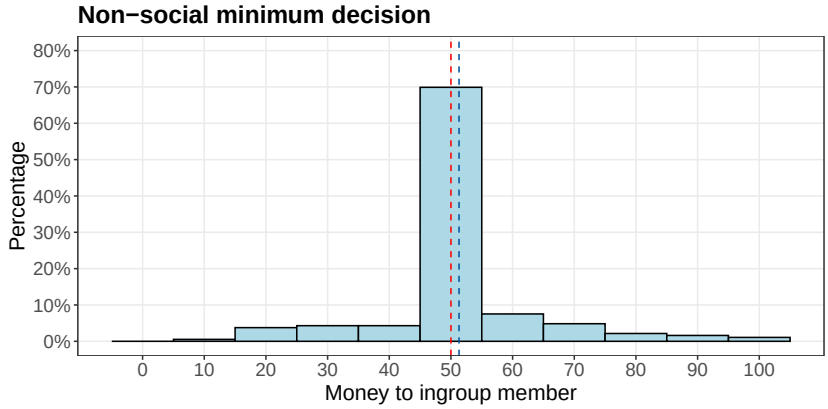
B Additional figures

Figure B.1: Incentive ingroup and Incentive outgroup robustness treatment results



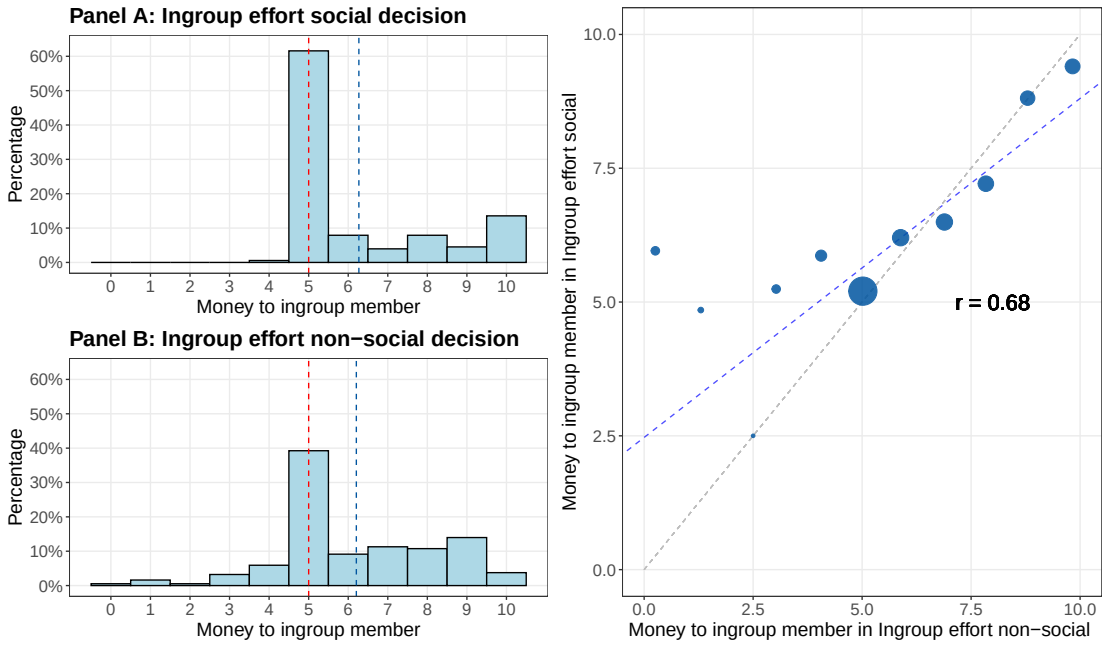
Notes: **Panel A and B:** Histogram of *Incentive ingroup* (Panel A) and *Incentive outgroup* (Panel B) decisions. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to the ingroup member instead of the outgroup member. Subjects incentive is to maximize the weighted sum of the in- and outgroup members WTP. In Panel A, the ingroup receives three times the weight, in Panel B, the outgroup receives three times the weight. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. For both panels, the binwidth is 10. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs), each panel thus displays $n = 180$ decisions by 60 subjects.

Figure B.2: Non-social minimum robustness treatment results



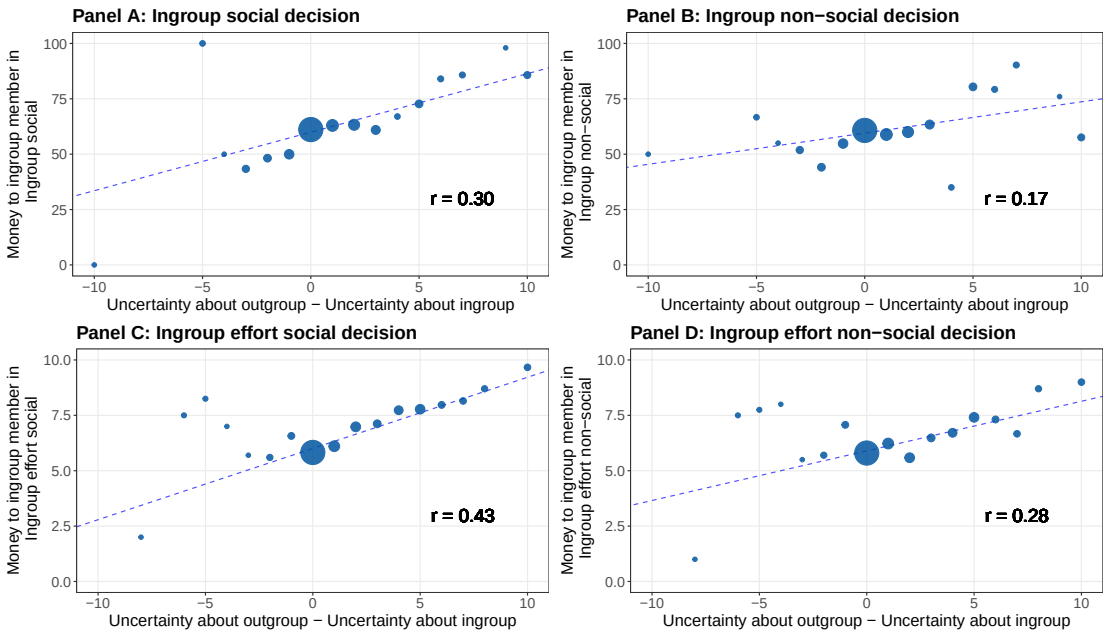
Notes: Histogram of *Non-social minimum* decisions. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to the ingroup member instead of the outgroup member. Subjects incentive is to maximize the minimum of the in- and outgroup member's WTP weighted with subjects' allocation. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. For both panels, the binwidth is 10. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs), each panel thus displays $n = 186$ decisions by 62 subjects.

Figure B.3: Results effort ingroup versus outgroup decisions



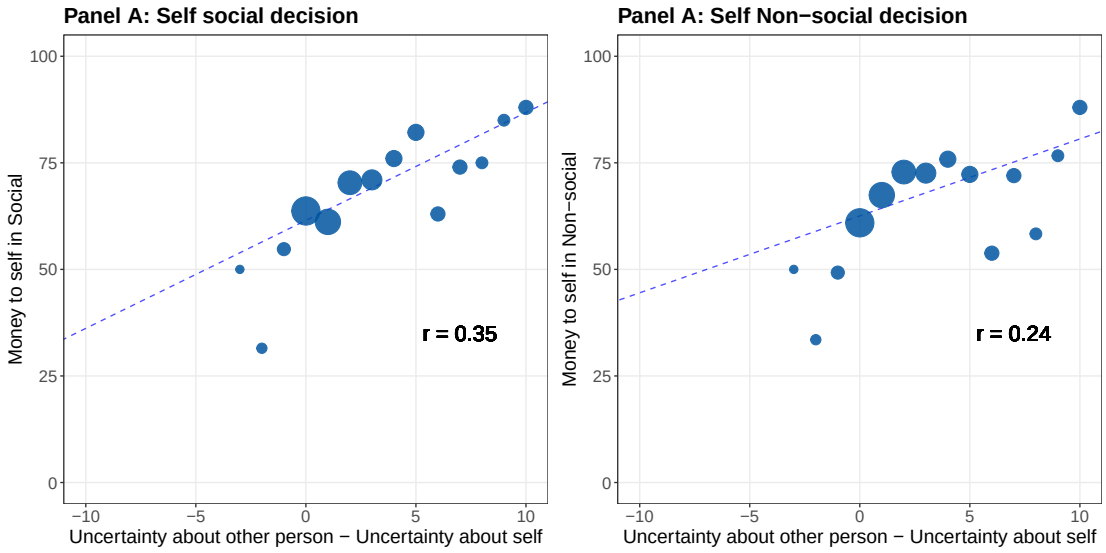
Notes: **Panel A and B:** Histogram of *Ingroup effort social* (Panel A) and *Ingroup effort non-social* (Panel B) decisions. The x-axis denotes the amount of money (out of \$10) that subjects allocate to the ingroup member instead of the outgroup member. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Ingroup social* (Panel A), the decisions have consequences for the ingroup and outgroup members. In *Ingroup non-social* (Panel B), the decisions have consequences only for the subjects, with their payoff depending on the ingroup and outgroup member's willingness-to-work on real effort tasks for the money. **Panel C:** Binned scatter plot of *Ingroup effort social* and *Ingroup effort non-social* decisions. The blue dotted line displays the linear fit of a regression of the *Ingroup effort social* on *Ingroup effort non-social* decisions. The correlation coefficient is $r = 0.53$. For all three panels, the binwidth is 1. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs), displaying $n = 177$ decisions by 59 subjects in Panel A, $n = 186$ decisions by 62 subjects in Panel B, and $n = 363$ decision-pairs by 121 subjects in Panel C.

Figure B.4: Association between interpersonal uncertainty and ingroup versus outgroup decisions



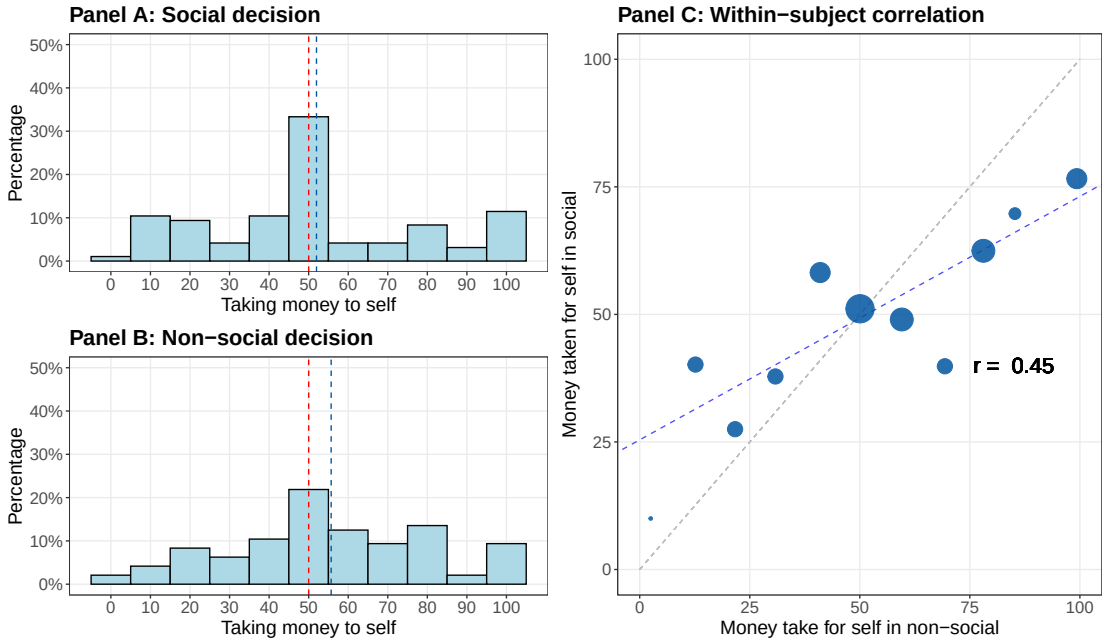
Notes: Binned scatter plot of the association of interpersonal uncertainty with *Ingroup social* decisions in Panel A, with *Ingroup non-social* decisions in Panel B, with *Ingroup effort social* decisions in Panel C, with *Ingroup effort non-social* decisions in Panel D. The x-axis denotes the difference in uncertainty ratings that subjects report about the ingroup and outgroup members gift card value. Higher values indicate more uncertainty about the outgroup's value relative to the ingroup. The y-axis denotes the amount of money that subjects allocate to the ingroup member instead of the outgroup member. In *Ingroup social* and *Ingroup effort social*, the decisions have consequences for the ingroup and outgroup members. In *Ingroup non-social* and *Ingroup effort non-social*, the decisions have consequences only for the subjects, with their payoff depending on the ingroup and outgroup member's (i) WTP for a \$100 gift card in case of *Ingroup non-social* and (ii) WTW for a \$10 bonus payment in case of *Ingroup effort non-social*. The blue dotted line displays the linear fit of a regression of the *social* decisions and *non-social* decisions, respectively, on the difference in interpersonal uncertainty ratings. The correlation coefficient is $r = 0.30$ in *Ingroup social*, $r = 0.17$ in *Ingroup non-social*, $r = 0.43$ in *Ingroup effort social*, and $r = 0.28$ in *Ingroup effort non-social*. The binwidth is 1. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs), displaying $n = 357$ decision-pairs by 119 subjects in Panels A and B and $n = 363$ decision-pairs by 121 subjects in Panels C and D.

Figure B.5: Association between interpersonal uncertainty and Dictator game decisions



Notes: Binned scatter plot of the association of interpersonal uncertainty with *Self social* decisions in Panel A and with *Self non-social* decisions in Panel A. The x-axis denotes the difference in uncertainty rating that subjects report about the other person's gift card value and about their own value. Higher values indicate more uncertainty about the other person's value relative to uncertainty about their own value. The y-axis denotes the amount of gift card money (out of $100) that subjects allocate to themselves instead of another individual. In *Self social* (Panel A), the decision has consequences for the subjects and the matched other individuals. In *Self non-social* (Panel B), the decision has consequences only for the subjects, with their payoff depending on their and the other individual's WTP for the gift card. The blue dotted line displays the linear fit of a regression of the *Self social* decisions and *Self Non-social* decisions, respectively, on the difference interpersonal uncertainty measure. The correlation coefficient is $r = 0.35$ in *Ingroup social* and $r = 0.24$ in *Ingroup social*. The binwidth is 1. Displayed are $n = 120$ decision-pairs by 120 subjects.

Figure B.6: Giving versus taking results excluding non-takers



Notes: **Panel A and B:** Histogram of the *Taking social* (Panel A) and *Taking non-social* (Panel B) decision. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to themselves instead of another individual. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Taking social* (Panel A), the other individual has earned the $100 and subjects decide how much to take away for themselves. Their decision has consequences for themselves and the other individual. In *Taking non-social* (Panel B), the decision has consequences only for the subjects, with their payoff depending on their WTP and the other individual's WTA for the gift card. **Panel C:** Binned scatter plot of the *Taking social* and *Taking non-social* decision. The blue dotted line displays the linear fit of a regression of the *Taking social* on the *Taking non-social* decision. The correlation coefficient is $r = 0.32$. For all three panels, the binwidth is 10. Excluded are subjects that take nothing from the other individual in *Taking social*. Thus, displayed are $n = 96$ decision-pairs by 96 subjects.

C Rawlsian preferences under interpersonal uncertainty

In Section 2, we showed that utilitarianism generates patterns of prosocial behavior given certain assumptions on interpersonal uncertainty. This raises the question of whether every commonly used welfare criterion delivers similar patterns under the right parameters given our assumptions. Here, we show that Rawlsian preferences – one of the most discussed welfare criterion – are insensitive to interpersonal uncertainty. Under Rawlsian preferences, only the utility of the least well-off recipient matters. In our context, Rawlsian preferences mean the utility individual i receives from allocating x to the ingroup member and $(100 - x)$ to the outgroup member is $u_{RAWLS} = \min\{v_1 x, v_2(100 - x)\}$. As Theorem 2 shows, a decision-maker will then split the money equally independent of differences in interpersonal uncertainty between recipients.

Theorem 2. *Suppose individual i has Rawlsian preferences. Then irrespective of i's risk attitude ($U'' \leq 0$ or $U'' \geq 0$), her optimal allocation is $x^* = 50$, in both the following cases, i) $f_1 = f_2$, and, ii) f_2 is a mean preserving spread of f_1.*

Proof of Theorem 2. As v_1, v_2 are random variables, i's expected utility from allocating x to the outgroup is:

$$EU(x) = E_{v_1 \sim f_1, v_2 \sim f_2} \min\{v_1 x, v_2(100 - x)\}$$

For any $x \in [0, 50) \cup (50, 100]$,

$$\begin{aligned} \min\{v_1(100 - x), v_2 x\} + \min\{v_1 x, v_2(100 - x)\} &\leq v_1(100 - x) + v_1 x \\ &= 100 v_1 \end{aligned}$$

with strict inequality whenever $v_1 \neq v_2$.[25]
Similarly, $\min\{v_1(100-x), v_2 x\} + \min\{v_1 x, v_2(100-x)\} \leq 100 v_2$ with strict inequality whenever $v_1 \neq v_2$. Putting these two inequalities together, we get

$$\min\{v_1(100 - x), v_2 x\} + \min\{v_1 x, v_2(100 - x)\} \leq \min\{100 v_1, 100 v_2\}$$

with strict inequality whenever $v_1 \neq v_2$. Next, using $f_1 = f_2$,

$$\begin{aligned} EU(x) &= E_{v_1 \sim f_1, v_2 \sim f_2} \min\{v_1 x, v_2(100 - x)\} \\ &= E_{v_1 \sim f_1, v_2 \sim f_2} \min\{v_1(100 - x), v_2 x\} \end{aligned}$$

[25] If $x = 50$, then strict inequality does not hold under $v_1 < v_2$.

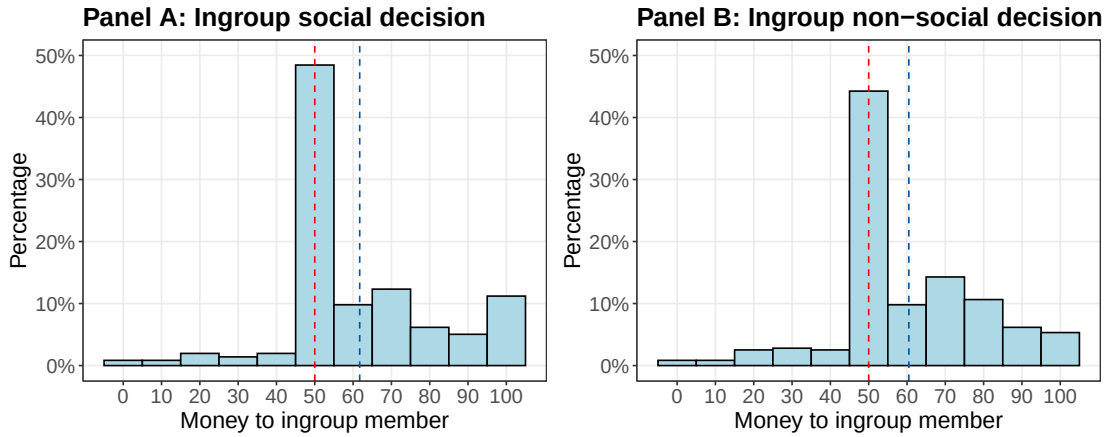
Therefore, for any $x \in [0, 50) \cup (50, 100]$,

$$\begin{aligned} EU(x) &= \frac{1}{2} \times E_{v_1 \sim f_1, v_2 \sim f_2} \left(\min\{v_1 x, v_2(100 - x) + \min\{v_1(100 - x), v_2 x\} \right) \\ &< \frac{1}{2} \times E_{v_1 \sim f_1, v_2 \sim f_2} \min\{100 v_1, 100 v_2\} \\ &= E_{v_1 \sim f_1, v_2 \sim f_2} \min\{50 v_1, 50 v_2\} \end{aligned}$$

The first inequality becomes strict as $v_1 \neq v_2$ with positive probability in the integration. This proves part (i), and a similar proof works for part (ii) after v_2 is replaced with $w_1 + z_1$ like in the proof of Theorem 1.

□

Figure D.1: Ingroup within-subject



Notes: Histogram of *Ingroup social* (Panel A) and *Ingroup non-social* (Panel B) decisions. The x-axis denotes the amount of gift card money (out of \$100) that subjects allocate to the ingroup member instead of the outgroup member. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Ingroup social* (Panel A), the decisions have consequences for the in- and outgroup members. In *Ingroup non-social* (Panel B), the decisions have consequences only for the subjects, with their payoff depending on the in- and outgroup member's WTP for the gift card. For both panels, the binwidth is 10. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs). Both panels display $n = 357$ decisions by 119 subjects.

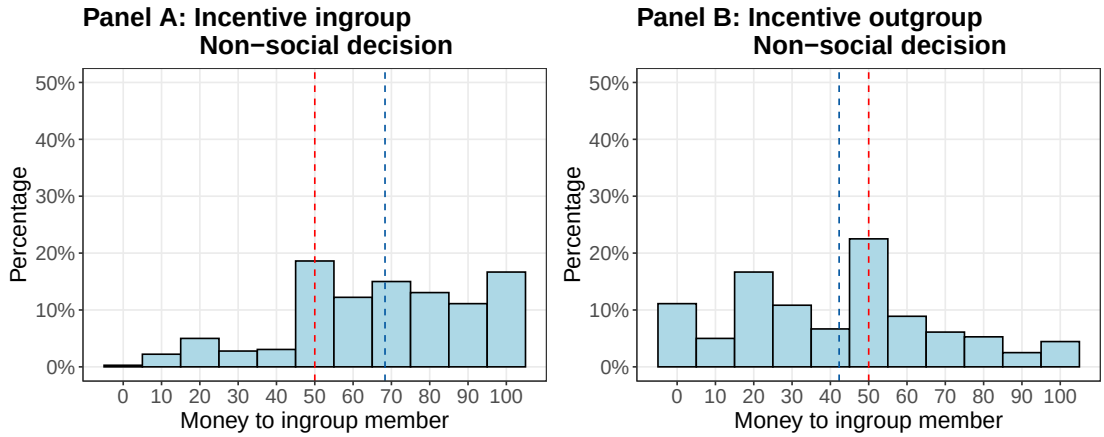
D Within-subject analyses

The results covered in the main text were obtained using a between-subject design, where we only used the first decision each subject faced. In the following, we repeat our analyses using all of the subjects' decisions. In general, our between-subject results replicate well in the within-subject analyses.

D.1 Ingroup versus outgroup paradigm main results

***Ingroup social* decisions.** In the within-subject case, subjects allocate on average \$57.48 if their ingroup members share the same interests/hobbies, \$67.81 if political views are shared, and \$59.88 if religious beliefs are shared. In all three cases, we can reject the hypothesis of no ingroup favoritism ($p < 0.001$, one-sample Wilcoxon tests). Figure D.1 panel A displays the distribution pooled over the three decisions. In 52% of the decisions, subjects display ingroup favoritism by allocating strictly more than 50% to the ingroup. Outgroup favoritism is found in 9% of decisions, and in the remaining 39%, subjects allocate 50/50. In total, 76% of subjects display ingroup favoritism in at least one decision.

Figure D.2: Ingroup incentive within-subject



Notes: Histogram of *Incentive ingroup* (Panel A) and *Incentive outgroup* (Panel B) decisions. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to the ingroup member instead of the outgroup member. Subjects' incentive is to maximize the weighted sum of the ingroup and outgroup members' WTP. In Panel A, the ingroup receives three times the weight, in Panel B, the outgroup receives three times the weight. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. For both panels, the binwidth is 10. Decisions are pooled across the three groups (shared hobbies/interests, political views, and religious beliefs). Each panel thus displays $n = 360$ decisions by 120 subjects.

***Ingroup non-social* decisions.** When facing the *Ingroup non-social* decisions first, subjects allocate on average $58.47 to their ingroup members sharing the same interests/hobbies, $64.00 if political views are shared an $58.97 if religious beliefs are shared. Again, we can reject the hypothesis of no ingroup favoritism ($p < 0.001$, one-sample Wilcoxon tests) in all three cases. Figure D.1 panel B displays the distribution. In 55% of the decisions, subjects display ingroup favoritism by allocating strictly more than 50% to the ingroup. Outgroup favoritism is found in 12% of decisions, and in the remaining 32%, subjects allocate 50/50.

Comparing *Ingroup social* and *non-social*. Comparing average ingroup allocations between *Ingroup social* and *non-social* between-subject reveals that we cannot reject equality in all three cases ($p = 0.59$ for hobbies/interests, $p = 0.20$ for political views, $p = 0.94$ for religious beliefs, paired Wilcoxon tests). Further, we cannot reject that the distributions of allocations are equal ($p = 0.30$ for hobbies/interests, $p = 0.23$ for political views, $p = 0.99$ for religious beliefs, Kolmogorov-Smirnov test). The same holds when decisions are pooled across the three domains for additional statistical power ($p = 0.40$, Kolmogorov-Smirnov test).

D.2 Ingroup versus outgroup setting incentive robustness

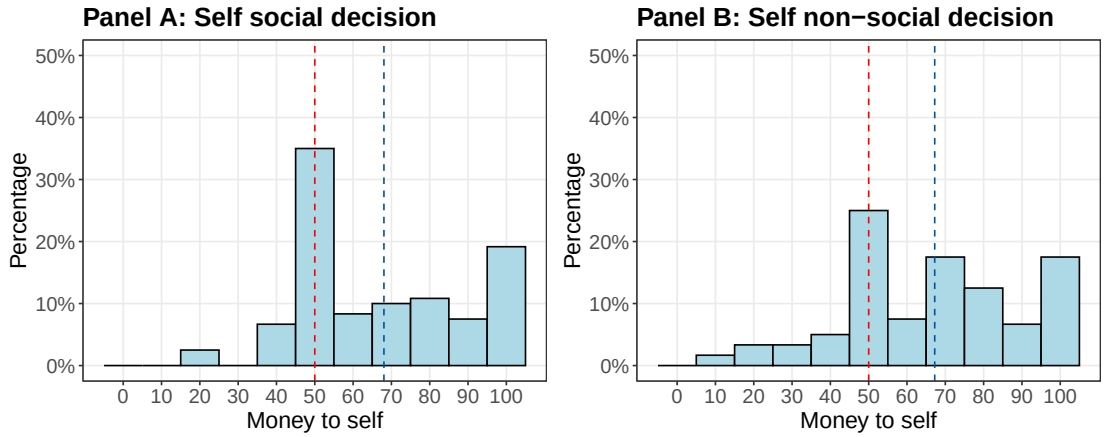
Table D.1 displays the treatment of *Outgroup incentive* relative to *Ingroup incentive* effects separately for the within-subject and between-subject effects pooled across the three groups. As displayed, the effect is similar in both the within- and between-subject comparison. Regarding the within-subject effects in the social groups individually, when the ingroup is incentivized, average ingroup allocations increase from \$58.47 to \$67.22 for hobbies/interests ($p < 0.001$, unpaired Wilcoxon tests), from \$64.00 to \$72.57 for political views ($p < 0.001$), and from \$58.97 to \$65.22 for religious beliefs ($p = 0.01$) compared to *Ingroup non-social*. Conversely, in *outgroup incentive*, allocations to the ingroup decrease to \$37.76 for hobbies/interests, to \$46.89 for politics and to \$42.21 for religious beliefs (all three $p < 0.001$). As in the between-subject comparison, we again see outgroup favoritism in the *Outgroup incentive* decisions. The pooled average is \$42.29, which is significantly different from the even split ($p < 0.001$, one-sample Wilcoxon test). See Figure D.2 for the distributions, which once again show that the shift in average giving is driven by shifts in the distributions.

Table D.1: Treatment effect of the incentive treatment in the ingroup versus outgroup setting

	Dependent variable:	
	Allocation to ingroup member	
	Within-subject (1)	Between-subject (2)
Outgroup incentive	-26.047^{***}	-25.817^{***}
	(3.163)	(3.547)
Constant (*Ingroup incentive*)	68.333^{***}	67.978^{***}
	(1.937)	(2.083)
Subjects	120	120
Observations	720	360
R^2	0.211	0.231

Notes: The table shows OLS estimates. The dependent variable in columns (1) and (2) is the amount subjects allocate to themselves (out of \$100) in the *Ingroup incentive* and *Ingroup incentive* treatments. "Outgroup incentive" is a dummy variable equal to one if the incentive for the decision gave three times the weight on the outgroup member's WTP, and equal to zero if the incentive gave three times the weight on the ingroup member's WTP. In column (1), all decisions are used, in (2) only the first decisions. Standard errors (in parentheses) are clustered at the subject level. Significance levels: *p<0.1, **p<0.05 and ***p<0.01.

Figure D.3: Dictator game between subject



Notes: Histogram of the *Self social* (Panel A) and *Self non-social* (Panel B) decision. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to themselves instead of another individual. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. In *Self social* (Panel A), the decision has consequences for the subjects and the other individual. In *Self non-social* (Panel B), the decision has consequences only for the subjects, with their payoff depending on their and the other individual's WTP for the gift card. For both panels, the bin-width is 10. Both panels display $n = 240$ decisions by 120 subjects.

D.3 Self versus others setting main results

***Self social* decision.** In the within-subject case of the *Self social* decision, subjects allocate on average $68.05 to themselves, thus displaying significant self-regarding behavior relative to the equal split ($p < 0.001$, one-sample Wilcoxon test). Figure D.3 panel A displays the distribution. In total, 62% of subjects allocate more money to themselves, 9% allocate more to the other person, and 29% implement the 50/50 split.

***Self non-social* decision.** When facing the *Self non-social* decision as the first decision, subjects allocate on average $67.02 to themselves, again displaying significant self-regarding behavior ($p < 0.001$, one-sample Wilcoxon tests). Figure D.3 panel B shows the distribution. In total, 66% of subjects allocate more money to themselves, 13% allocate more to the other person, and 21% implement the 50/50 split.

Comparing *Self social* and *non-social*. In the between-subject comparison, we also cannot reject equality of average allocations between *Self social* and *non-social* ($p = 0.69$, paired Wilcoxon tests). Similarly, we cannot reject that the pooled distributions are equal ($p = 0.95$, Kolmogorov-Smirnov test).

E Analyzing order effects

A potential concern for the validity of the within-subject results is contagion across conditions. As subjects facing the first set of decisions were not aware that a second set would follow, this naturally cannot influence our between-subject analysis presented in the main paper that only uses the first set of decisions. However, subjects may adjust their choice in the subsequent *non-social* decisions to mimic the *social* decisions, potentially biasing the individual-level analyses. Such adjustment could lead to artificially high similarity between the two decisions, and thus artificially high correlations. Because we randomized the order of decisions, we can directly assess this concern by testing for order effects. Overall, we find no evidence that the order influences subjects' behavior, as we show in the following in detail.

E.1 Ingroup versus outgroup paradigm

For the *Ingroup non-social* decisions, the pooled average allocations to the ingroup are \$60.89 when elicited before, and \$60.10 when elicited after the *social* decisions. For hobbies/interests *Ingroup non-social* the averages are \$56.86 and \$59.95 ($p = 0.58$, unpaired Wilcoxon test), for political views \$65.02 and \$63.06 ($p = 0.61$), and for religious beliefs \$60.81 and \$57.29 ($p = 0.15$). Thus, the averages are invariant to the order. We also cannot reject the null that distributions are invariant to the order ($p = 0.61, p = 0.31, p = 0.22$, Kolmogorov-Smirnov tests). Moving to the *Ingroup social* decisions, average allocations are \$63.38 when *Ingroup social* is elicited first, and \$59.92 when elicited after the *non-social* decisions. Again, averages and distributions generally do not differ significantly. For hobbies/interests *Ingroup social* the averages are \$57.48 and \$57.47 ($p = 0.54$, unpaired Wilcoxon test), for political views \$71.05 and \$64.28 ($p = 0.26$), and for religious beliefs \$61.61 and \$58.00 ($p = 0.99$). We also cannot reject the null that distributions are invariant to the order for hobbies/interests and religious beliefs ($p = 0.27$ and $p = 0.32$, Kolmogorov-Smirnov tests), with the only exception being political views ($p = 0.01$).

E.2 Self versus other paradigm

In the case of the *Self non-social* decisions, subjects allocate \$64.12 to themselves when the decision is before the *Self social* decision, and\$70.25 when the decision comes afterward, an insignificant difference ($p = 0.12$, unpaired Wilcoxon test). In the case of the *Self social* decisions, subjects allocate \$69.05 to themselves when the decision is before the *Self non-social* decision, and \$67.02 when the decision comes

afterward, again an insignificant difference ($p = 0.61$). In addition, we can reject the null that the distributions are invariant to the order at the 5% level for*Self non-social* ($p = 0.08$ Kolmogorov-Smirnov tests) and at any conventional level for *Self social* ($p = 0.54$ Kolmogorov-Smirnov tests).

E.3 Giving versus taking paradigm

In the case of the *Taking non-social* decisions, subjects allocate $56.00 to themselves when the decision is before the *Taking social* decision, and $54.05 when the decision comes afterward, an insignificant difference ($p = 0.49$, unpaired Wilcoxon test). In the case of the *Taking social* decisions, subjects allocate $39.52 to themselves when the decision is before the *Taking non-social* decision, and $41.51 when the decision comes afterward, again an insignificant difference ($p = 0.52$). In addition, we cannot reject the null that the distributions are invariant to the order both for *Taking non-social* ($p = 0.45$ Kolmogorov-Smirnov tests) and *Taking social* ($p = 0.73$ Kolmogorov-Smirnov tests).

F Self versus others setting incentive treatment

Design. As in the ingroup case, we vary the incentive subjects face when making the *Self non-social* decisions. In *Self incentive*, the weight on the DM's own WTP is three times as high as the other individuals WTP. The DM's payoff thus becomes:

$$\Pi\left(x_{self}, x_{other}\right) = 3 \cdot x_{self} \cdot WTP_{self}/100 + x_{other} \cdot WTP_{other}/100$$

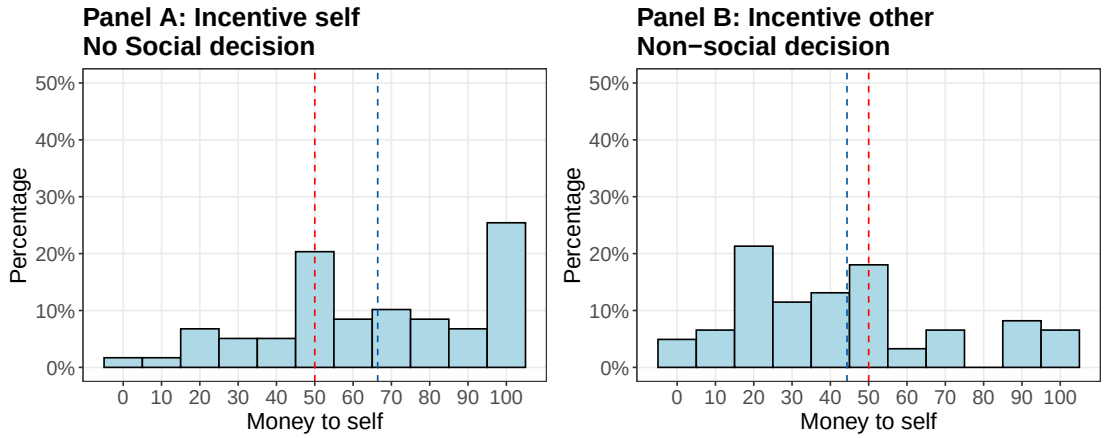
In *Other incentive* we increase the weight put on the other individual's WTP to be three times as high as the DM's WTP:

$$\Pi\left(x_{self}, x_{other}\right) = x_{self} \cdot WTP_{self}/100 + 3 \cdot x_{other} \cdot WTP_{other}/100$$

Results. Inducing these incentives changes people's behavior in the *non-social* decision. See Table F.1 for the within-subject and between-subject treatment effect. In both cases lead the change in incentives to a significant change in the amount subjects allocate to themselves, they allocate \$19.50 in the within and \$22.08 in the between-subject comparison less to the themselves when the incentives are higher for the other participant. Figure F.1 displays the distributions in the between-subject case. The fraction of subjects allocating more than 50% of the endowment to themselves increases from 33% in *Other incentive* to 63% in *Self incentive*, while the fraction of subjects allocating more money to the other participant decreases from 50% to 22%. The distributions are significantly different from each other ($p < 0.001$, Kolmogorov-Smirnov test.

Using the within-subject comparison shows that 83% of subjects change their allocation behavior between *Self incentive* and *Other incentive*. Among those 17% subjects that are unresponsive to the incentive change, 33% allocate more to themselves, a substantially lower fraction than the 58% in the main *Self non-social* case. In total, 38% choose the equal split, and 29% allocate more to the other participant. Taking the behavior of these unresponsive subjects as indicative of inattention or confusion, it appears that such factors are associate with subjects allocating less to themselves. This result thus provides suggestive evidence that our replication of significantly more self-giving in *Self social* using the *Self non-social* decisions is not driven by inattentive or confused subjects.

Figure F.1: Self versus other incentive



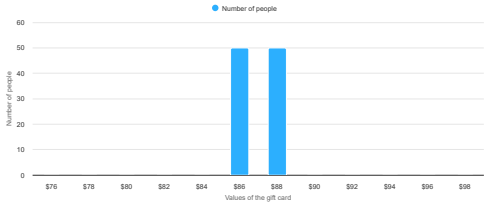
Notes: Histogram of *Self incentive* (Panel A) and *Other incentive* (Panel B) decisions. The x-axis denotes the amount of gift card money (out of $100) that subjects allocate to themselves instead of another individual. Subjects incentive is to maximize the weighted sum of their own and another individuals WTP. In Panel A, subjects own WTP receives three times the weight, in Panel B, the other individual's WTP receives three times the weight. The red dotted line denotes the even split benchmark, the blue dotted line the average allocation. For both panels, the binwidth is 10. Only the first decision is used for each subject. Panel A displays $n = 59$ decisions by 59 subjects, Panel B displays $n = 61$ decisions by 61 subjects.

Table F.1: Treatment effect of the incentive treatment in the self versus other setting

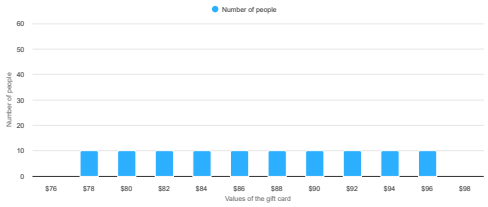
	Dependent variable:	
	Allocation to self	
	Within-subject (1)	Between-subject (2)
Other incentive	-19.500^{***}	-22.079^{***}
	(3.468)	(5.045)
Constant (*Self incentive*)	65.625^{***}	66.424^{***}
	(2.563)	(3.592)
Subjects	120	120
Observations	240	120
R^2	0.108	0.140

Notes: The table shows OLS estimates. The dependent variable in columns (1) and (2) is the amount subjects allocate to themselves (out of $100) in the *Other incentive* and *Self incentive* treatments. "Other incentive" is a dummy variable equal to one if the incentive for the decision gave three times the weight on the other person's WTP, and equal to zero if the incentive gave three times the weight on the subject's own WTP. In column (1), all decisions are used, in (2) only the first decisions. Standard errors (in parentheses) are clustered at the subject level. Significance levels: *p$<$0.1, **p$<$0.05 and ***p$<$0.01.

Figure G.1: Distributions shown to subjects



(a) Distribution with low variance



(b) Distribution with high variance

G Validation of uncertainty measure

Our self-reported interpersonal uncertainty measure is intended to proxy whether subjects perceive higher interpersonal uncertainty of one group over another, as defined in Definition 1 in Section 2. However, it could be the case that subjects instead only report their perception about mean differences or concepts unrelated to uncertainty. In this section, we validate that our measure is indeed sensitive to those changes in interpersonal uncertainty captured by our definition. To do so, we provide subjects with two objective WTP distributions, one being a mean-preserving spread of the other, and investigate the impact on the answers subjects give to our measure.

Design. At the end of the *Ingroup belief measurement* and *Self/other belief measurement* treatments, we showed subjects two figures. Each figure displayed a frequency distribution of the WTP values of 100 fictitious individuals. In one, 50 individuals had a WTP of \$86 and 50 a WTP of \$88 (low variance distribution). In the other were 10 individuals for each of the 10 values between \$78 and \$96 (high variance distribution). See Figure G.1 for the figures shown to subjects. We also provided these values to subjects in text format below the figures. The high variance distribution is a mean-preserving spread of the low variance one, having the same mean but a lower variance. For each figure, subjects were asked the following about the group displayed in the figure:

> "Suppose we randomly pick one of the 100 people from this group. How certain are you about how much the randomly chosen person would value the Amazon gift card money?"

Subjects could respond on an 11-point Likert scale from *Very uncertain* to *Very certain*, and we re-code the variable so that higher values indicate higher uncertainty. The text and measurement thus closely mirror our self-reported interpersonal uncertainty measure.

Results. We find that subjects report different uncertainty across the two distributions. On average, they report an uncertainty of 3.46 Likert-scale points for the low variance distribution, and an uncertainty of 5.80 points for the high variance distribution, a significant difference ($p < 0.001$, paired Wilcoxon-test). On the individual level, 74% of subjects report a higher uncertainty for the high variance distribution compared to the low variance distribution, 14% report no difference, and the remaining 13% report more uncertainty for the low variance distribution. Thus, subjects are sensitive to changes in WTP distributions in the expected direction.

H Research transparency

All experiments covered in the paper were preregistered at aspredicted.org. The pre-registrations include details on the experimental design, the planned sample size, exclusion criteria, hypotheses, and the main analyses. Table H.1 provides an overview over the treatments and links to the respective pre-registrations.

Table H.1: Overview over treatments

Label	N	Covered in	Description
Ingroup social & *Ingroup non-social*	119	Section 3	Pre-reg link: https://aspredicted.org/H81_KQ5
Ingroup incentive & *Outgroup incentive*	120	Section 3.3	Pre-reg link: https://aspredicted.org/H81_KQ5
Ingroup minimum	62	Section 3.3	Pre-reg link: https://aspredicted.org/J7H_W8R
Ingroup effort social & *Ingroup effort non-social*	121	Section 3.3	Pre-reg link: https://aspredicted.org/53G_PNJ
Ingroup uncertainty	120	Section 3.4	Pre-reg link: https://aspredicted.org/53G_PNJ
Self social & *Self non-social*	120	Section 4	Pre-reg link: https://aspredicted.org/ZMF_CD9
Self incentive & *Other incentive*	120	App. Section F	Pre-reg link: https://aspredicted.org/ZMF_CD9
Self taking social & *Self taking non-social*	123	Section 5	Pre-reg link: https://aspredicted.org/RT4_TQB
Ingroup belief measurement & *Self/other belief measurement*	120	Sections 3.2, 4.2, 5.2	Pre-reg link: https://aspredicted.org/T7X_747

Our experimental implementation followed closely the pre-registration. In particular, we implemented the experimental design and sample size exactly as specified in the pre-registration. Similarly, we employed the exclusion criteria as pre-registered: we specified to exclude any subject who did not complete the experiment. This lead to the exclusion of 22 subjects in the *Ingroup social* and *Ingroup non-social* treatments, 28 in the *Ingroup incentive* and *Outgroup incentive*, 6 in the *Ingroup minimum*, 13 in the *Ingroup effort social* and *Ingroup effort non-social*, 15 in *Ingroup uncertainty*, 22 in *Self social* and *Self non-social*, 23 in *Self incentive* and *Other incentive*, 25 in *Self taking social* and *Self taking non-social*, 3 in *Ingroup belief measurement* and 1 in *Self/other belief measurement*. The sample sizes reported in Table H.1 are the final sample sizes used in all analyses of the paper after excluding the previously mentioned numbers of subjects.

H.1 Mapping the pre-registered hypotheses to the reported results

In the following, we map our pre-registered hypotheses to the results reported in the paper. In general, we find support for our pre-registered hypotheses.

Treatments *Ingroup social* and *Ingroup non-social*. Our first hypothesis was that in *Ingroup non-social*, subjects allocate more money to the ingroup than to the outgroup, i.e., that allocations are different from the 50/50 benchmark. Our second hypothesis was that there is a positive within-subject correlation of allocation behavior between the *Ingroup social* and *Ingroup non-social* decisions. Both hypotheses are supported by the data, the underlying results are described in Section 3.2. We further pre-registered as exploratory analysis the association of the perceived interpersonal uncertainty measure and decision behavior. This analysis is reported in Section 3.4.1.

Treatments *Ingroup incentive* and *Outgroup incentive*. Our hypothesis was that there would be more allocations to the outgroup member in *Outgroup incentive* than in *Ingroup incentive*. This hypothesis is supported by the data, the underlying results are described in Section 3.3.1.

Treatment *Ingroup minimum*. We hypothesized that there would be more allocations to the outgroup member in *Ingroup minimum* compared to *Ingroup non-social*. This hypothesis is supported by the data, the underlying results are described in Section 3.3.2.

Treatments *Ingroup effort social* and *Ingroup effort non-social*. Our first hypothesis was that in *Ingroup effort non-social*, subjects allocate more money to the ingroup than to the outgroup, i.e., that allocations are different from the 50/50 benchmark. Our second hypothesis was that there is a positive within-subject correlation of allocation behavior between the *Ingroup effort social* and *Ingroup effort non-social* decisions. Both hypotheses are supported by the data, the underlying results are described in Section 3.3.3. We further pre-registered as exploratory analysis the association of the perceived interpersonal uncertainty measure and decision behavior. This analysis is reported in Section 3.4.1.

Treatment *Ingroup uncertainty*. Our hypothesis was that subjects would be averse to variation in the willingness-to-work within groups. This hypothesis is supported by the data, the underlying results are described in Section 3.4.2. We further mentioned as exploratory analysis the use of a structural model, the model is reported in Section 3.5.2. The details of the model were not preregistered.

Treatments *Self social* and *Self non-social*. Our first hypothesis was that in *Self non-social*, subjects allocate more money to themselves than to the other individual, i.e., that allocations are different from the 50/50 benchmark. Our second hypothesis was that there is a positive within-subject correlation of allocation behavior between the *Self social* and *Self non-social* decisions. Both hypotheses are supported by the data, the underlying results are described in Section 4.2. We further pre-registered as exploratory analysis the association of the perceived interpersonal uncertainty measure and decision behavior. This analysis is reported in Section 4.3.

Treatments *Self incentive* and *Other incentive*. Our hypothesis was that there would be more allocations to the other participant in *Self incentive* than in *Other incentive*. This hypothesis is supported by the data, the underlying results are described in Appendix Section F.

Treatments *Self taking social* and *Self taking non-social*. Our first two hypotheses were that there would be more allocations to the other participant in *Self taking social* than in *Self social* and more in *Self taking non-social* than in *Self non-social*. Our third hypothesis was that there is a positive within-subject correlation of allocation behavior between the *Self taking social* and *Self taking non-social* decisions. All three hypotheses are supported by the data, the underlying results are described in Section 5.2. We further pre-registered as secondary hypothesis that subjects believe that people's willingness to accept is higher than their willingness to pay for gift card money. This hypothesis is supported by the data, the underlying result is described in 5.1.

H.2 Deviations from the pre-registration.

The pre-registrations for the *Ingroup social* and *Ingroup non-social* treatments as well as the *Self social* and *Self non-social* treatments contain another set of treatments labeled *Group info* and *Self info*. These treatments are not part of this paper and their results are available upon request because the design is superseded by the *Ingroup uncertainty* experiment.[26] The analyses contained in Section 3.5 were pre-registered as exploratory analyses without specific details.

[26]The omitted treatments show that providing subjects with information on the WTP of the recipients significantly changes their allocation behavior both in *Ingroup social* and *Self social*. However, in contrast to the *Ingroup uncertainty* experiment, this information manipulation does not directly manipulate interpersonal uncertainty and is potentially confounded by experimenter demand effects.

I Experimental instructions

This section provides screenshots of the experimental instructions. Note that in order to avoid anchoring effects, the slider-thumbs are initially hidden and only appear once subjects click on the slider-scale. For illustrative purposes, in some screenshots we clicked on the scale prior to making the screenshot.

I.1 Ingroup versus outgroup paradigm

I.1.1 Ingroup social and non-social screens

Figure I.1: Ingroup social and non-social screen 1

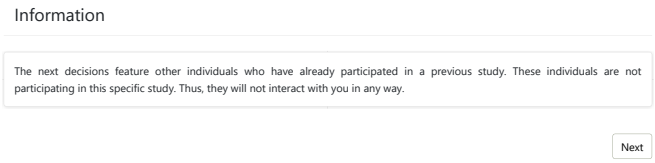


Figure I.2: Ingroup social and non-social screen 2

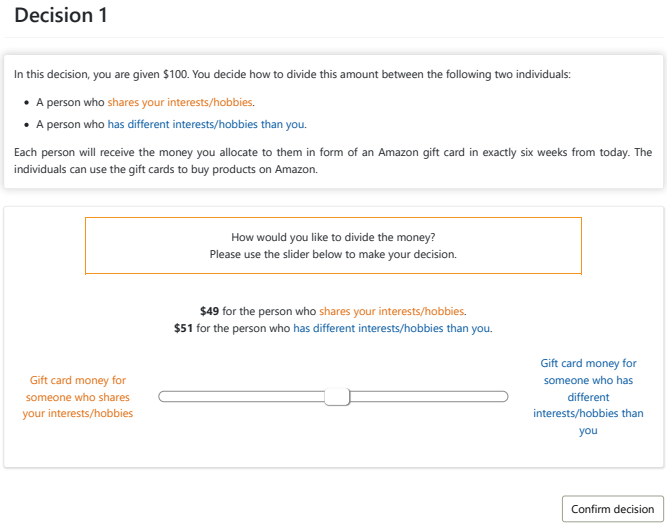


Figure I.3: Ingroup social and non-social screen 3

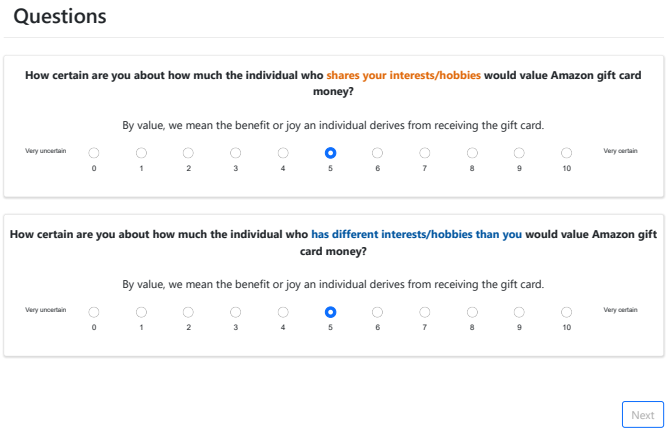


Figure I.4: Ingroup social and non-social screen 4

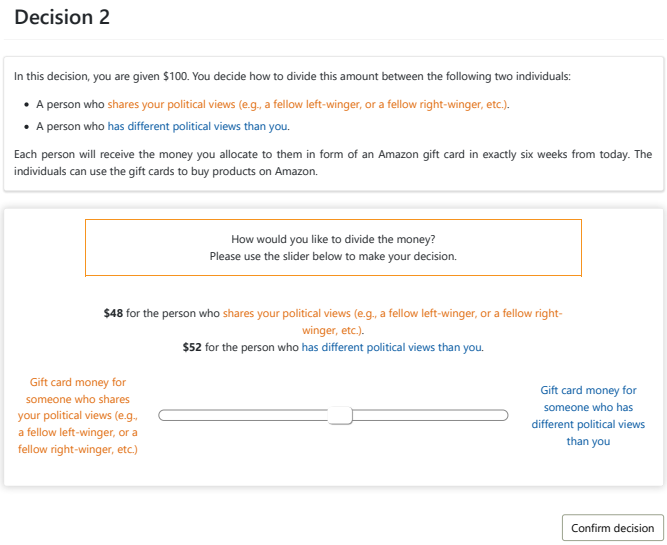


Figure I.5: Ingroup social and non-social screen 5

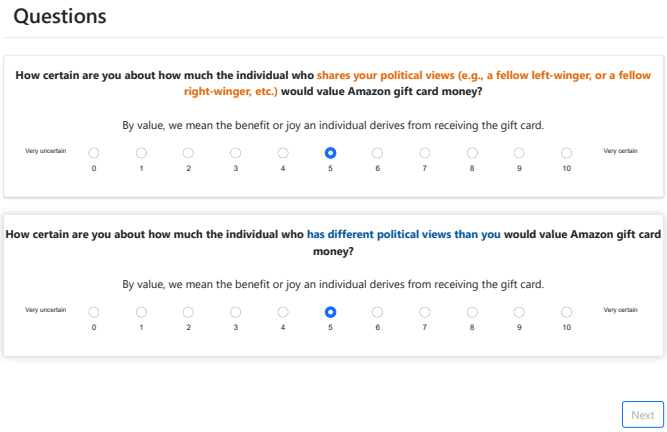


Figure I.6: Ingroup social and non-social screen 6

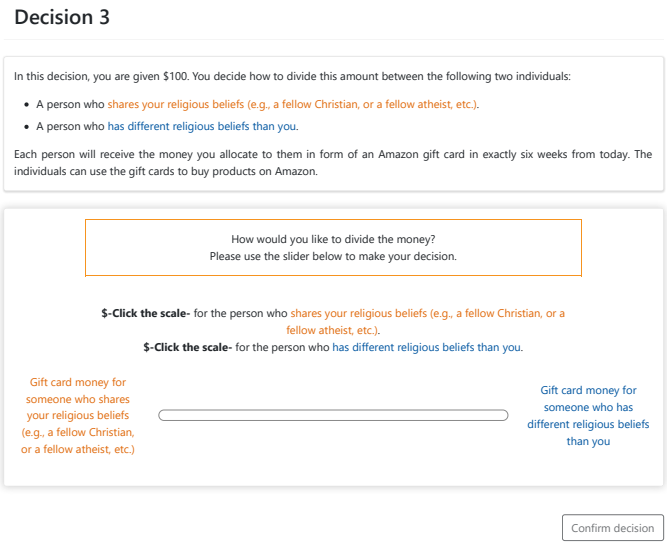


Figure I.7: Ingroup social and non-social screen 7

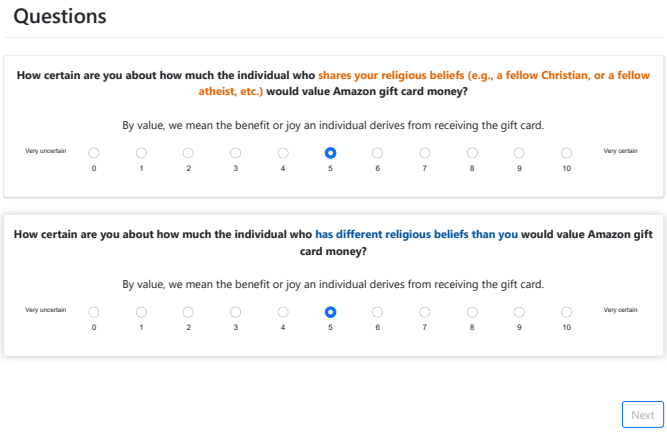


Figure I.8: Ingroup social and non-social screen 8

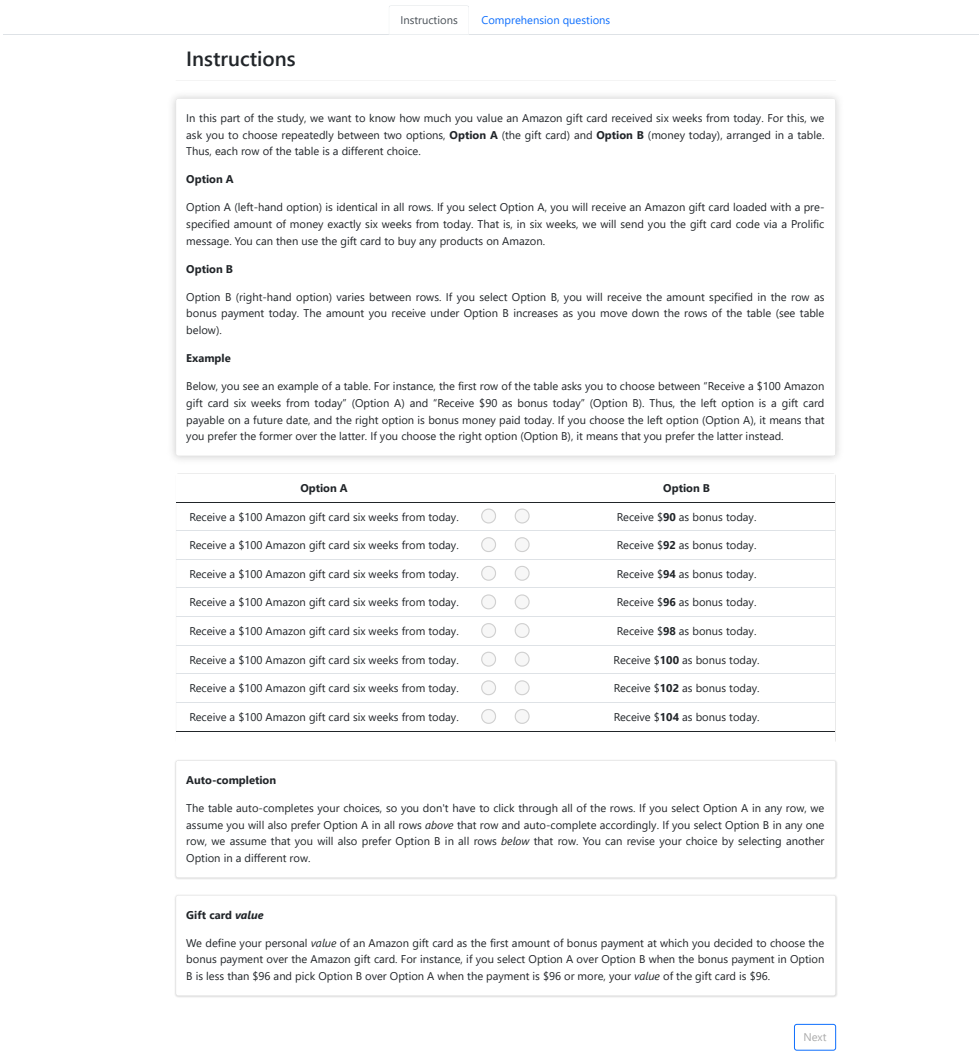


Figure I.9: Ingroup social and non-social screen 9

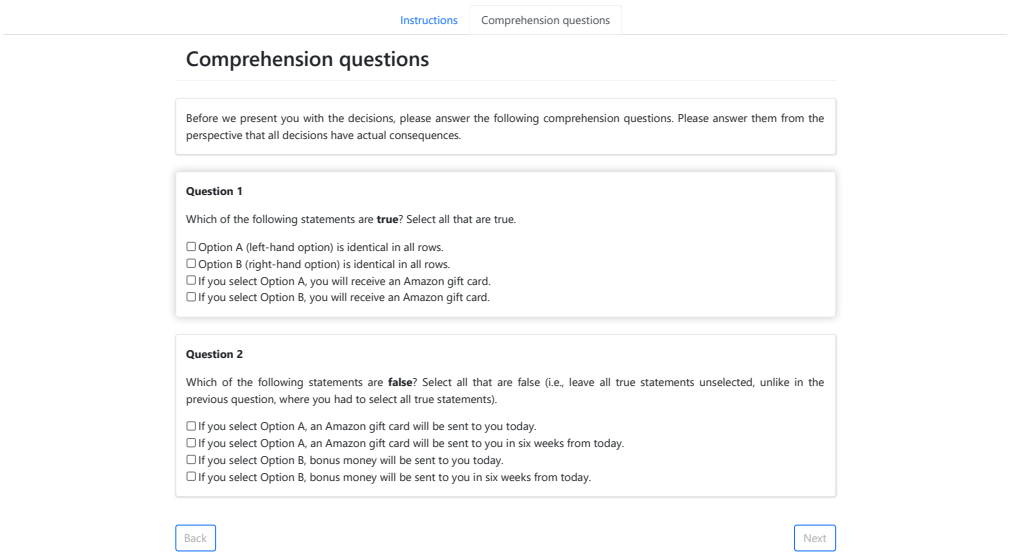


Figure I.10: Ingroup social and non-social screen 10

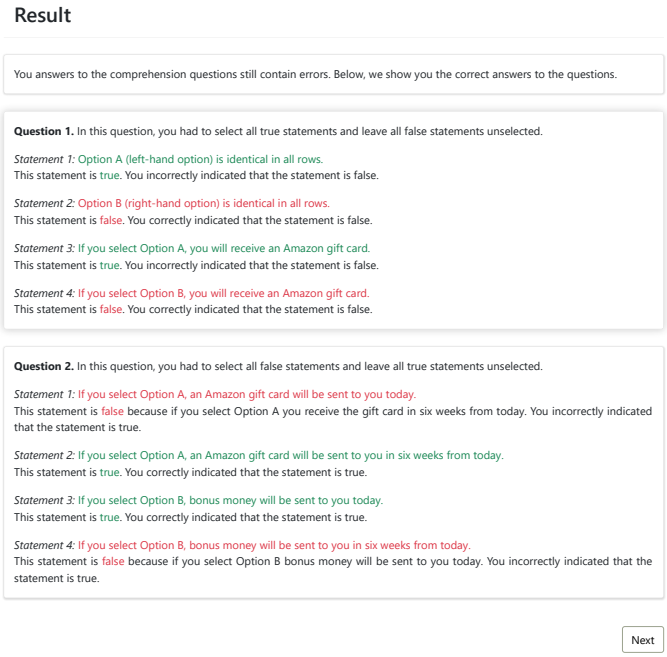


Figure I.11: Ingroup social and non-social screen 11

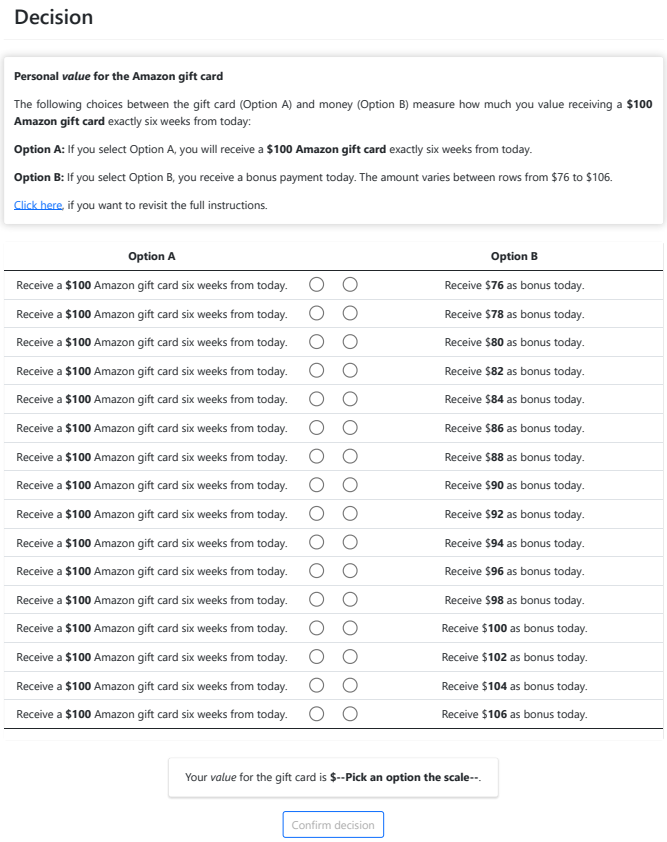


Figure I.12: Ingroup social and non-social screen 12

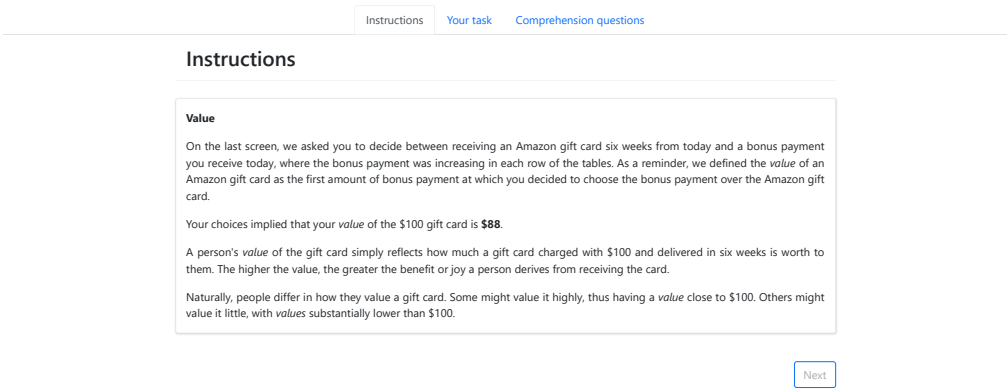


Figure I.13: Ingroup social and non-social screen 13

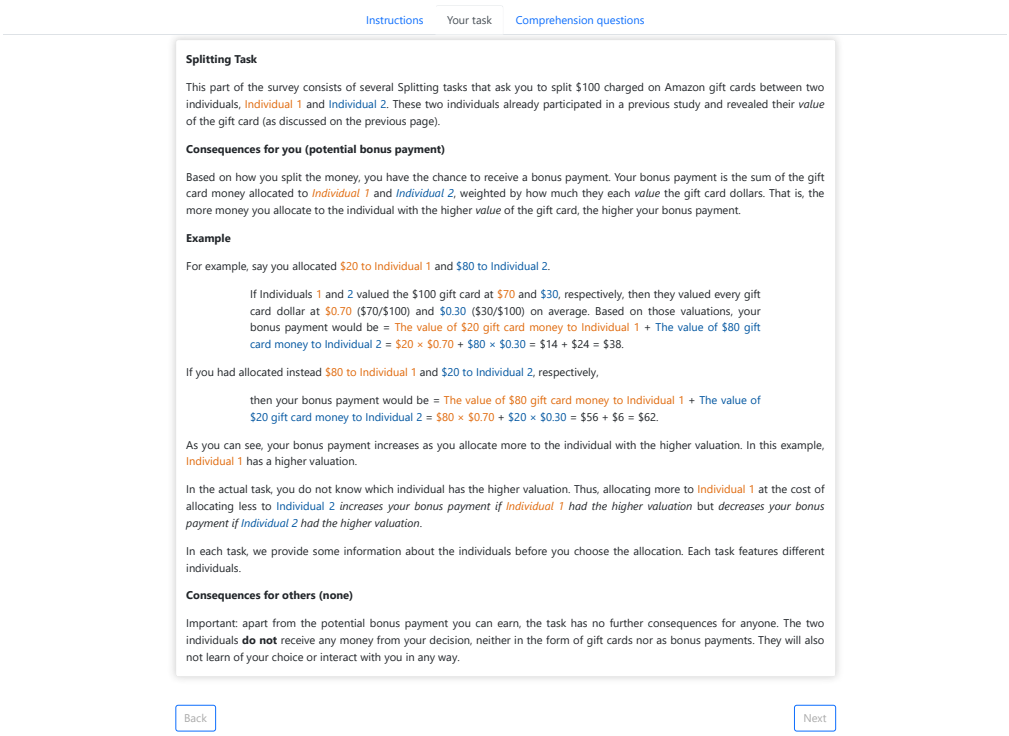


Figure I.14: Ingroup social and non-social screen 14

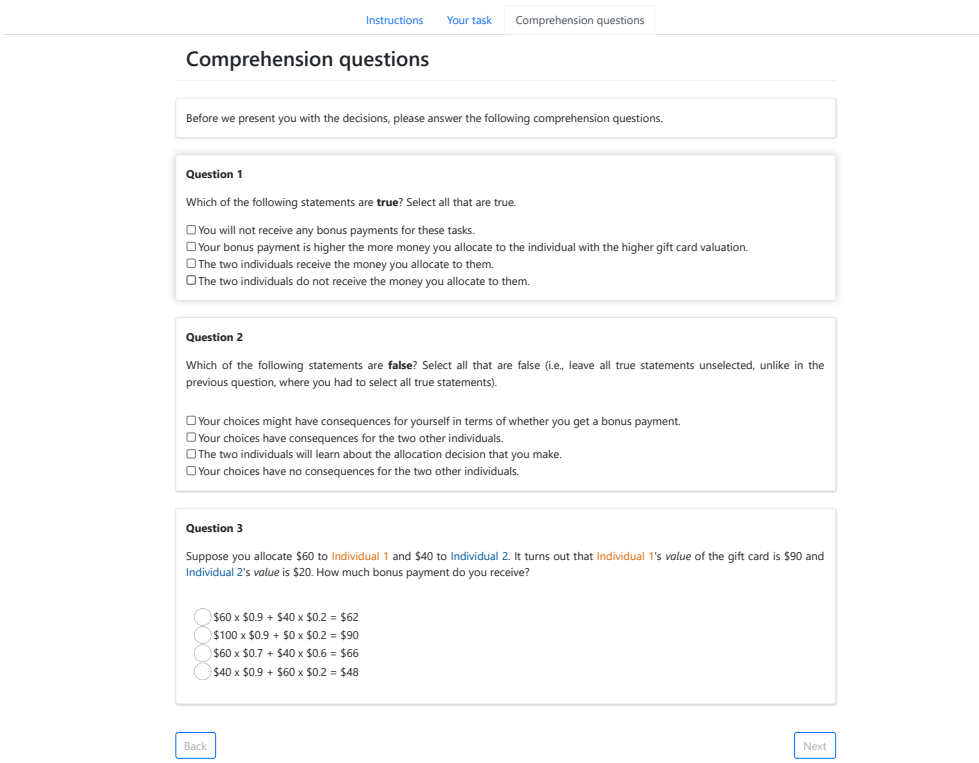


Figure I.15: Ingroup social and non-social screen 15

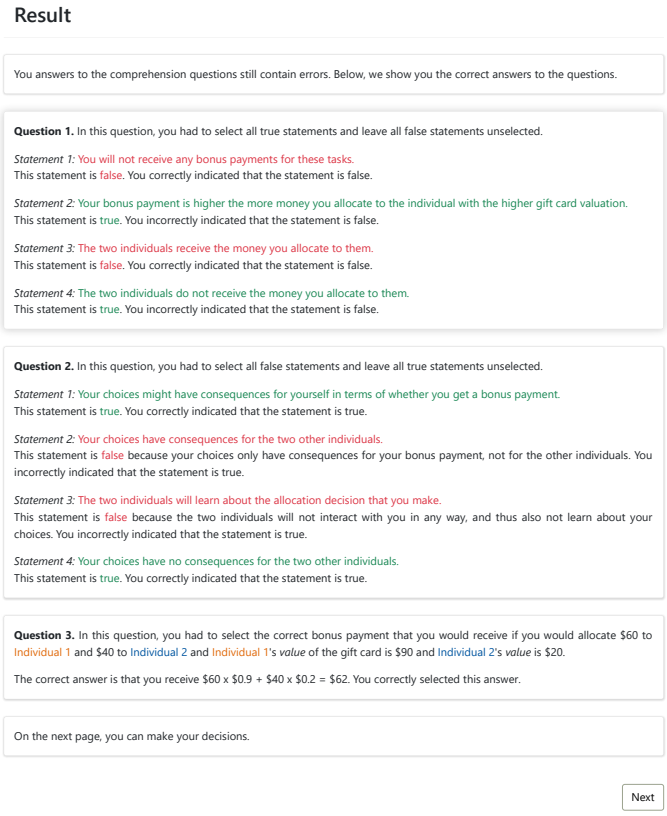


Figure I.16: Ingroup social and non-social screen 16

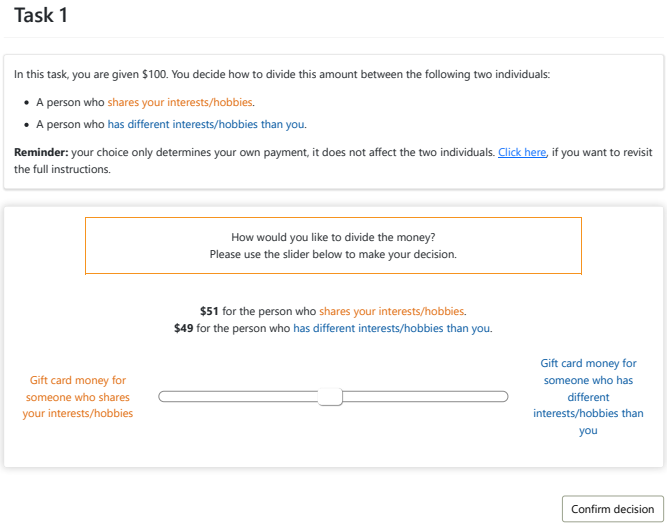


Figure I.17: Ingroup social and non-social screen 17

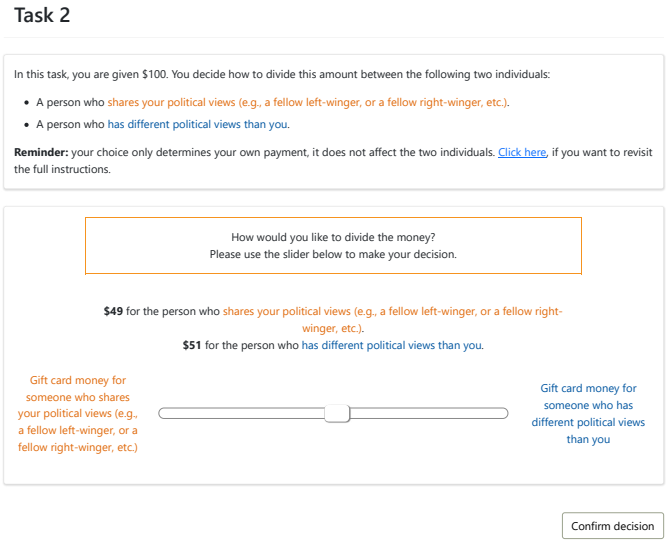
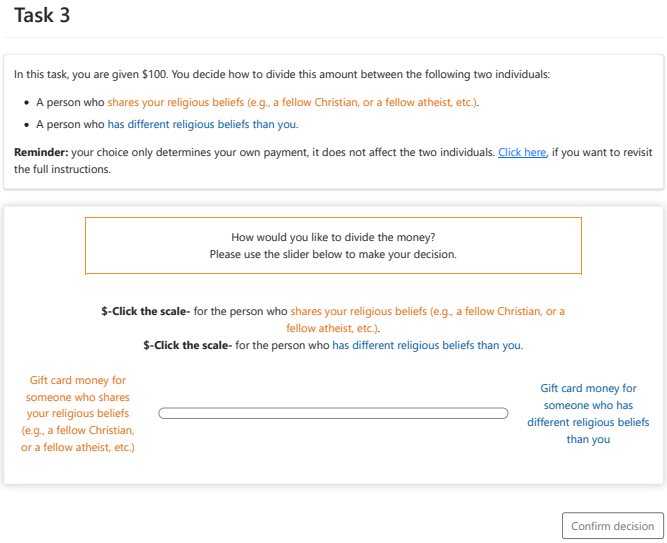


Figure I.18: Ingroup social and non-social screen 18



I.1.2 Incentive ingroup and Incentive outgroup screens

Figure I.19: Incentive ingroup and Incentive outgroup screen 1

Figure I.20: Incentive ingroup and Incentive outgroup screen 2

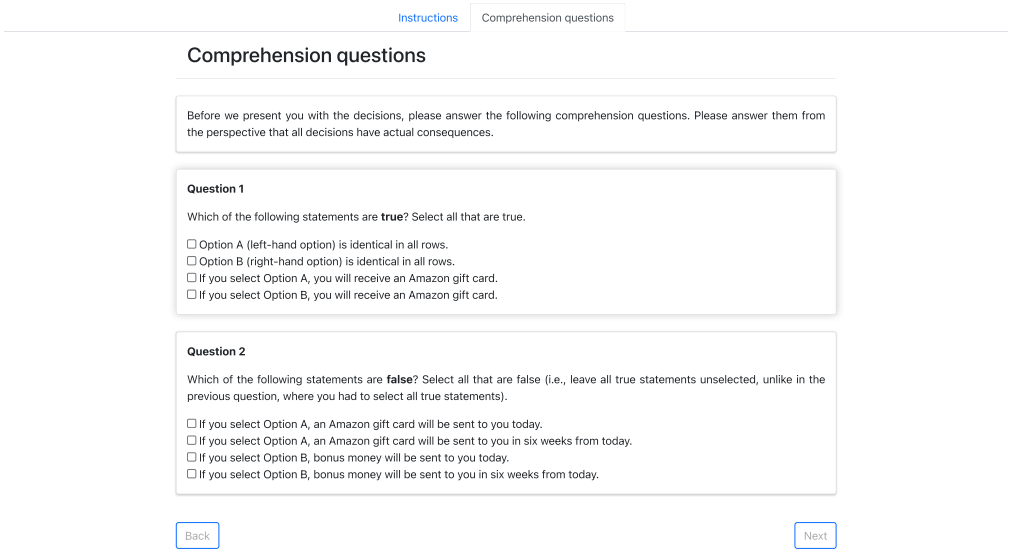


Figure I.21: Incentive ingroup and Incentive outgroup screen 3

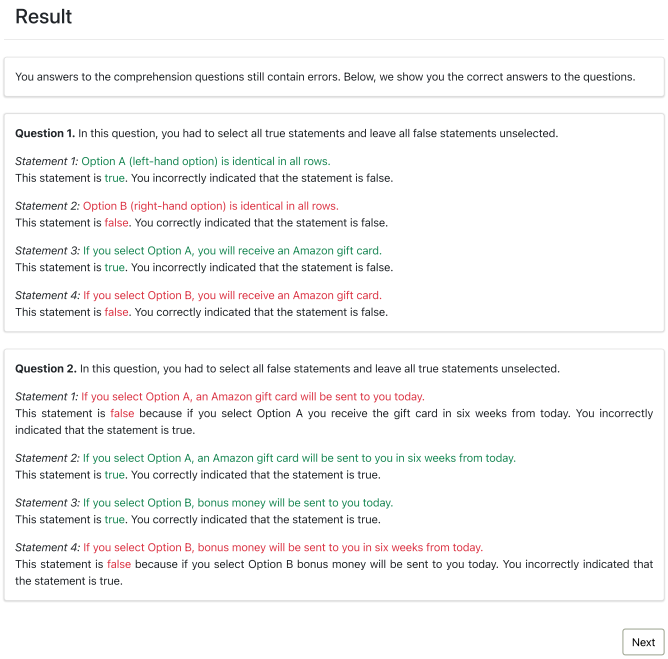


Figure I.22: Incentive ingroup and Incentive outgroup screen 4

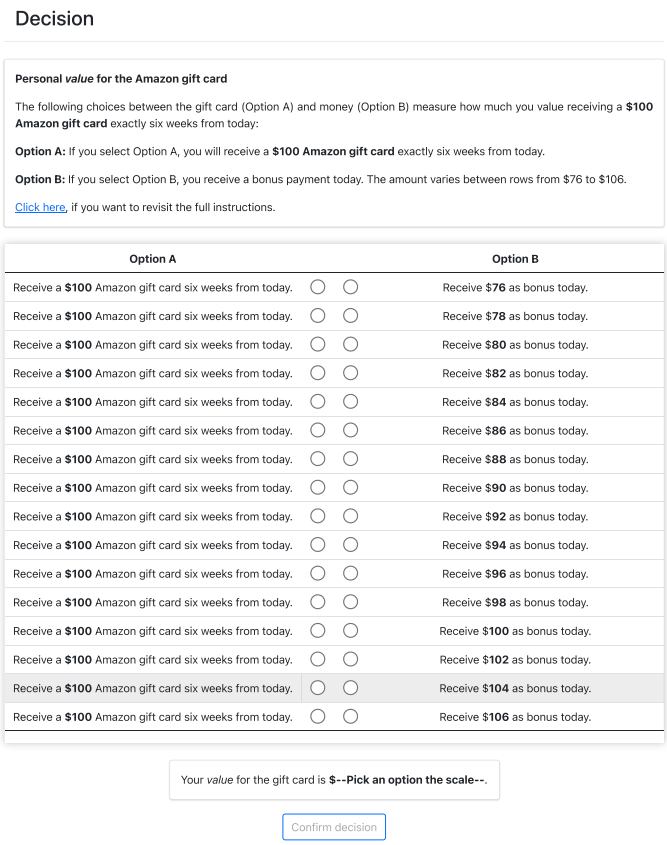


Figure I.23: Incentive ingroup and Incentive outgroup screen 5

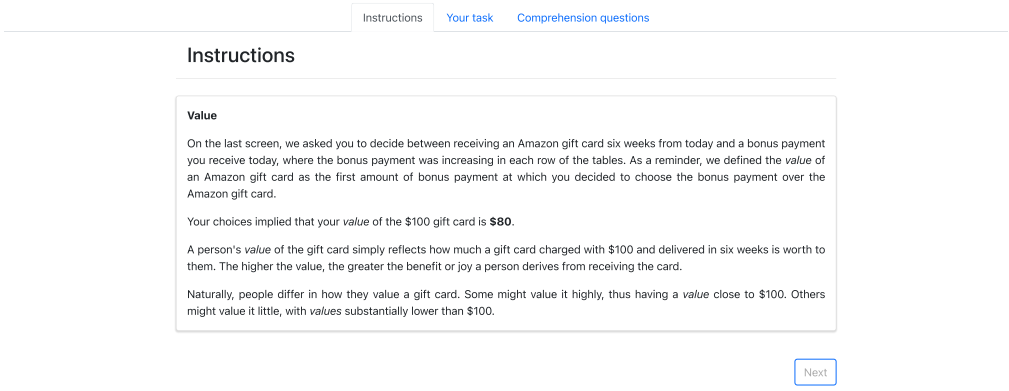


Figure I.24: Incentive ingroup and Incentive outgroup screen 6

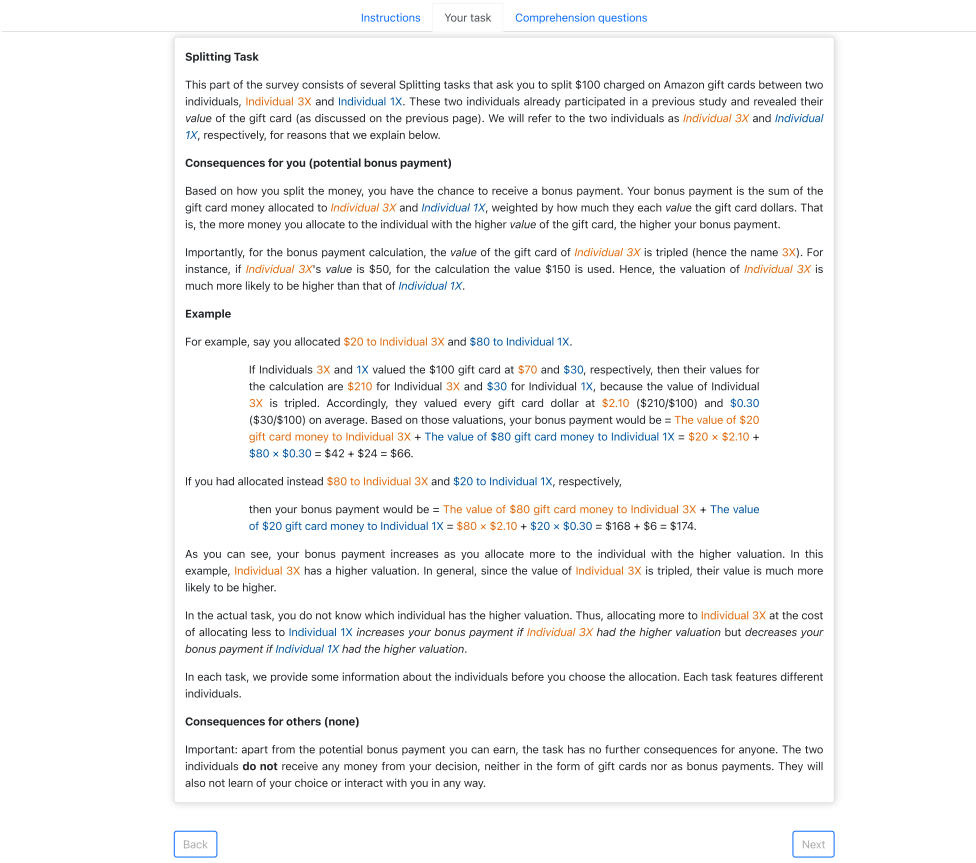


Figure I.25: Incentive ingroup and Incentive outgroup screen 7

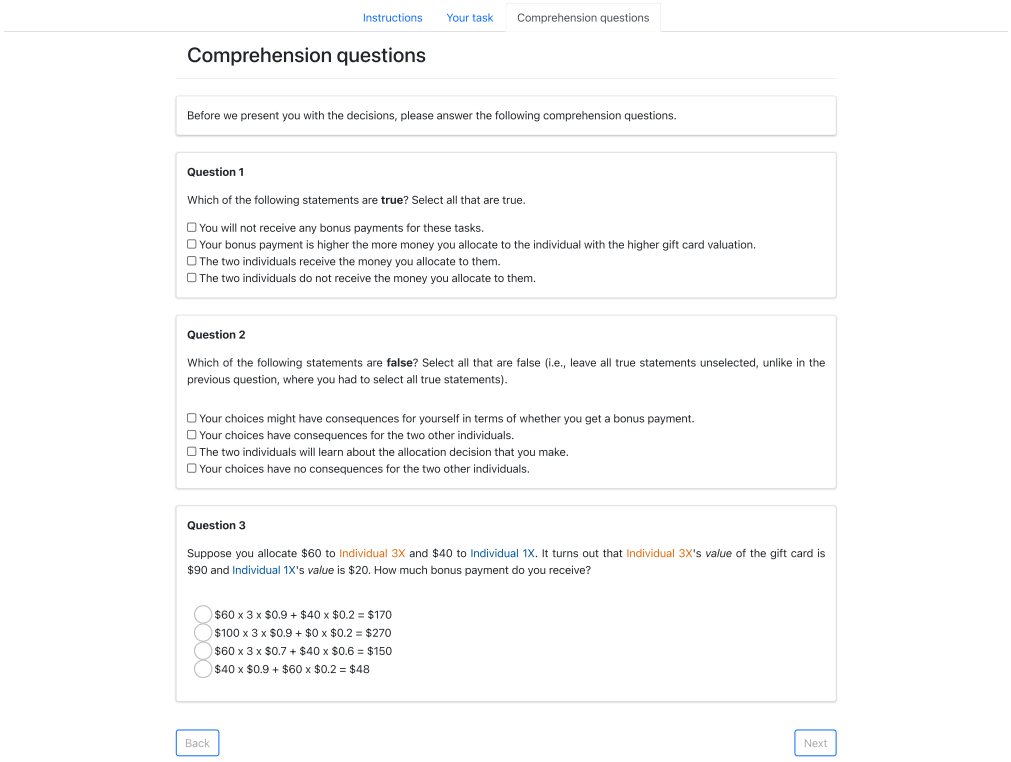


Figure I.26: Incentive ingroup and Incentive outgroup screen 8

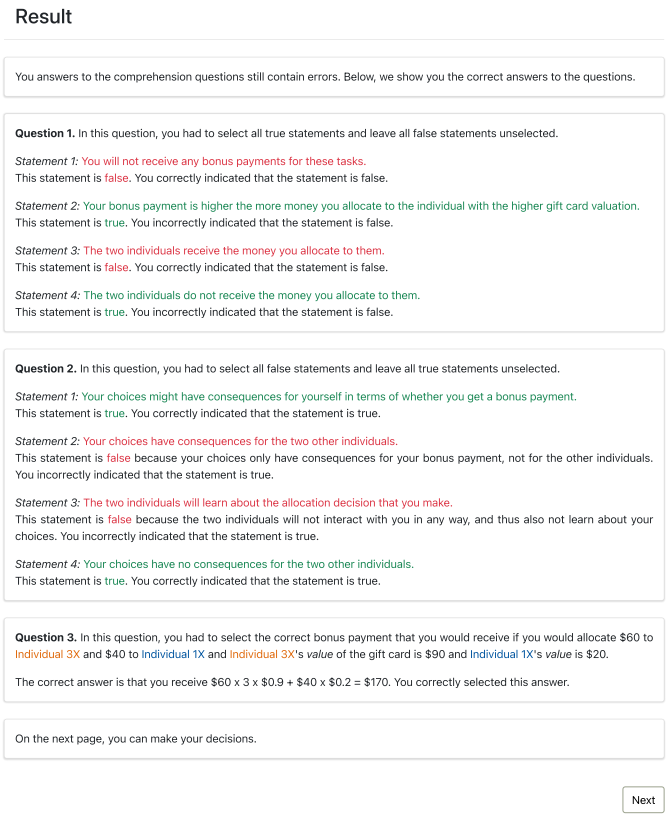


Figure I.27: Incentive ingroup and Incentive outgroup screen 9

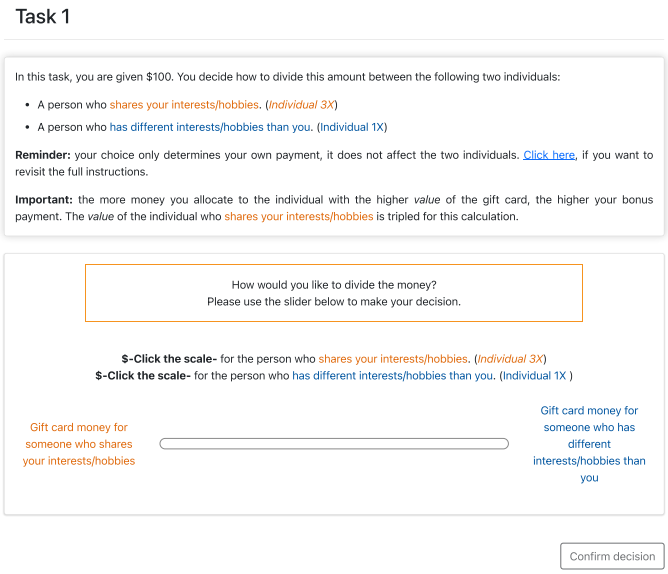


Figure I.28: Incentive ingroup and Incentive outgroup screen 10

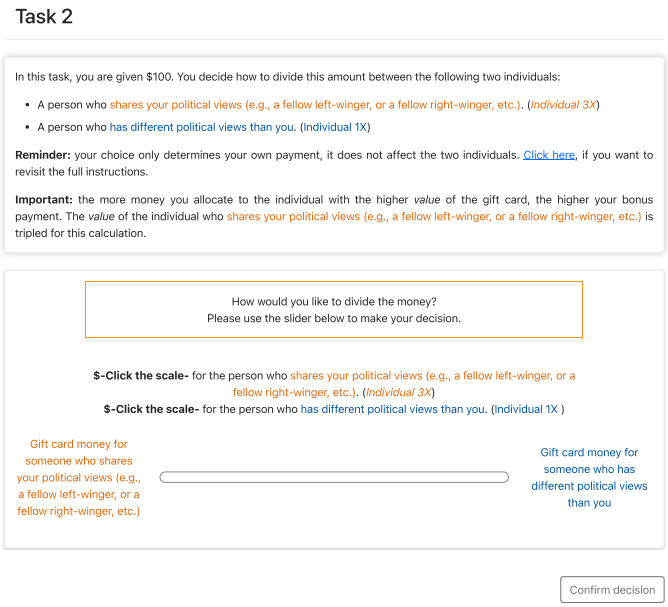


Figure I.29: Incentive ingroup and Incentive outgroup screen 11

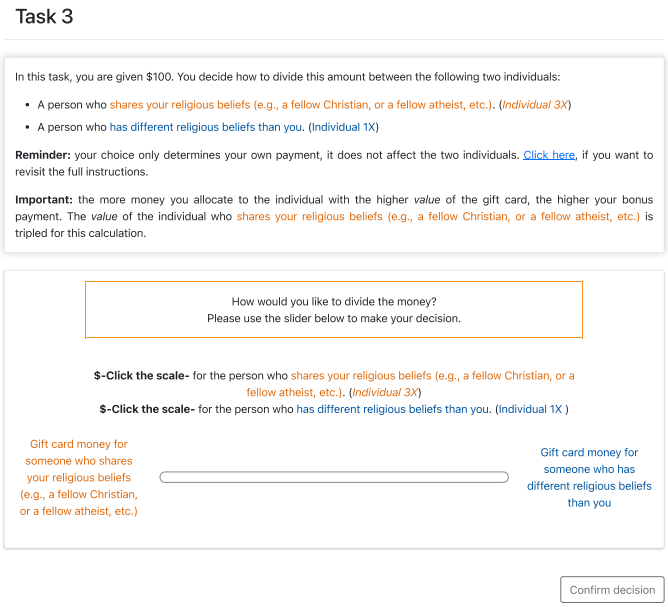


Figure I.30: Incentive ingroup and Incentive outgroup screen 12

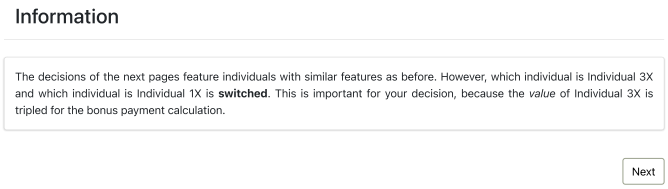


Figure I.31: Incentive ingroup and Incentive outgroup screen 13

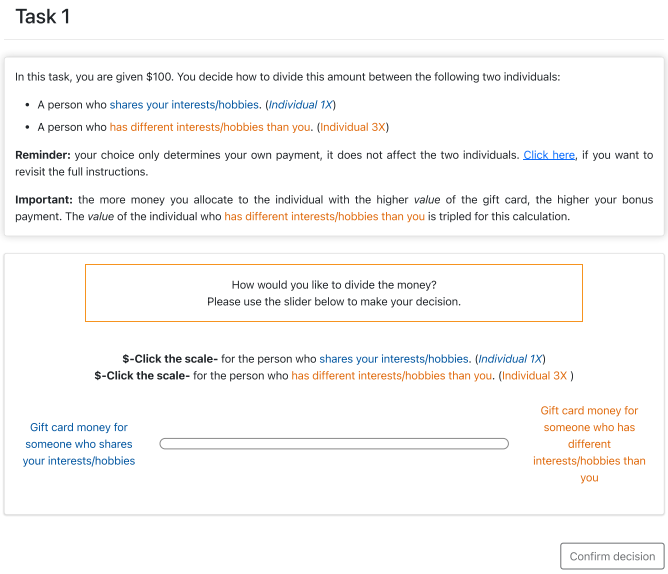


Figure I.32: Incentive ingroup and Incentive outgroup screen 14

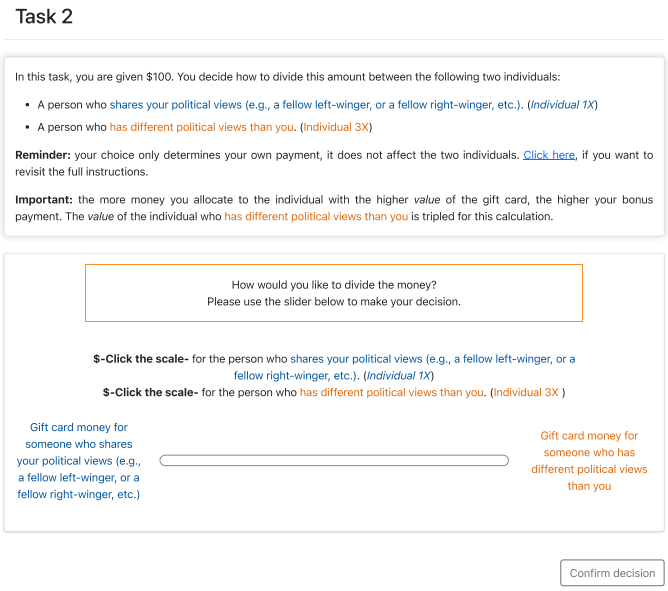
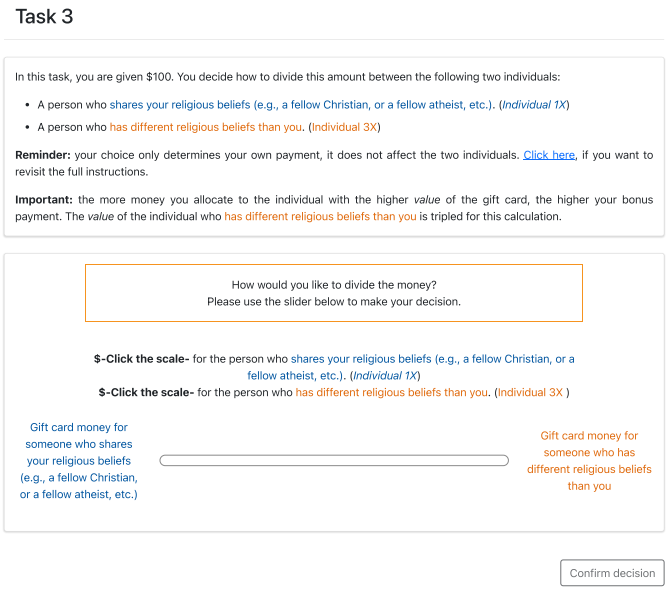


Figure I.33: Incentive ingroup and Incentive outgroup screen 15



I.1.3 Ingroup non-social minimum screens

Figure I.34: Ingroup non-social minimum screen 1

Figure I.35: Ingroup non-social minimum screen 2

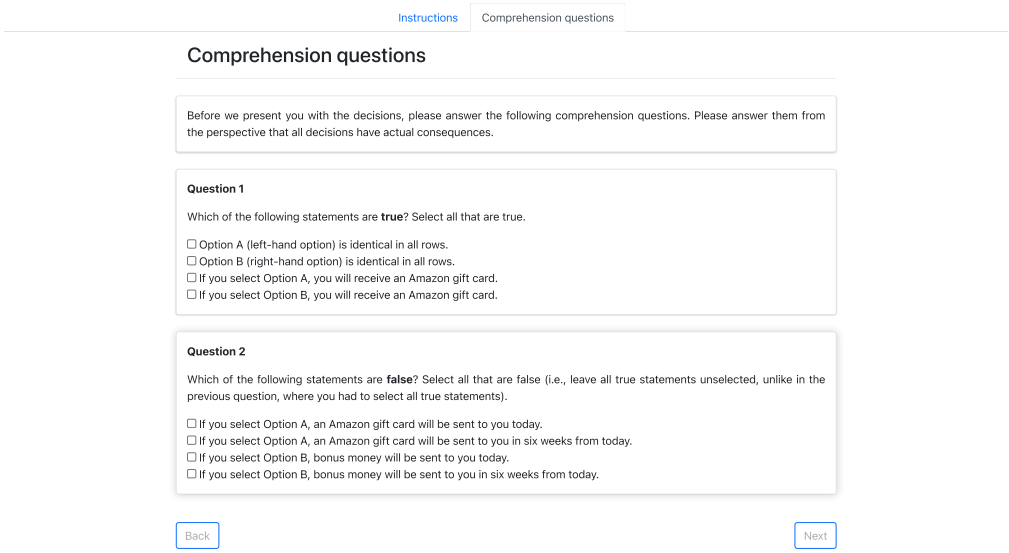


Figure I.36: Ingroup non-social minimum screen 3

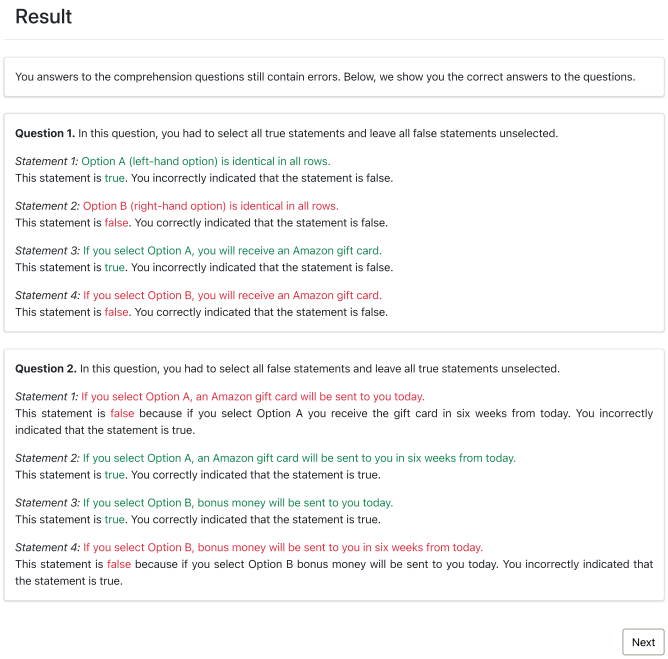


Figure I.37: Ingroup non-social minimum screen 4

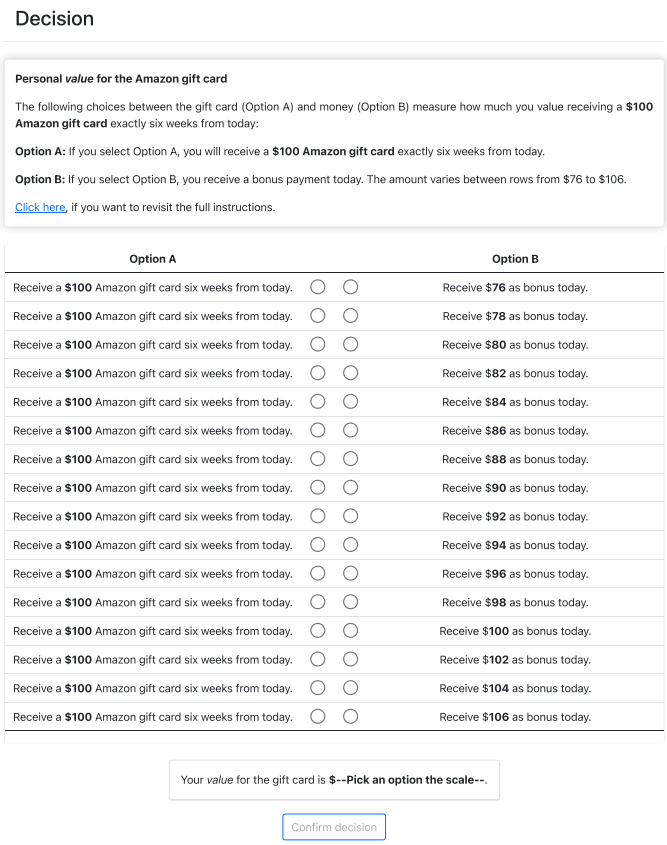


Figure I.38: Ingroup non-social minimum screen 5

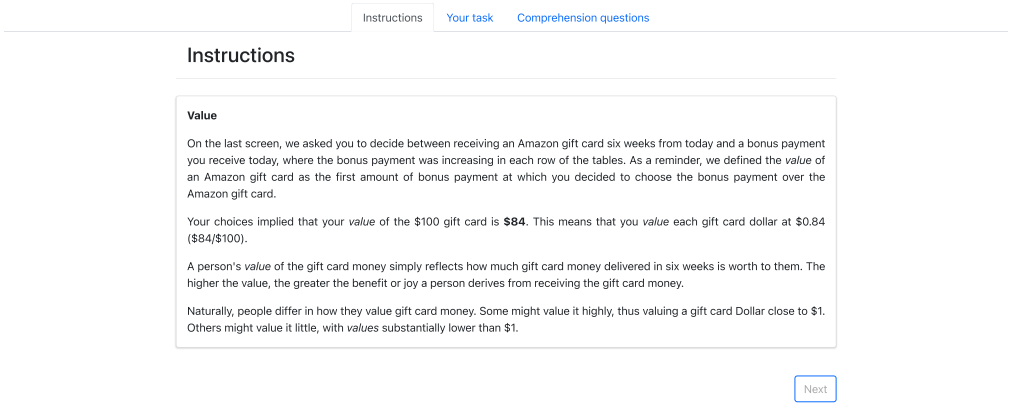


Figure I.39: Ingroup non-social minimum screen 6

Figure I.40: Ingroup non-social minimum screen 7

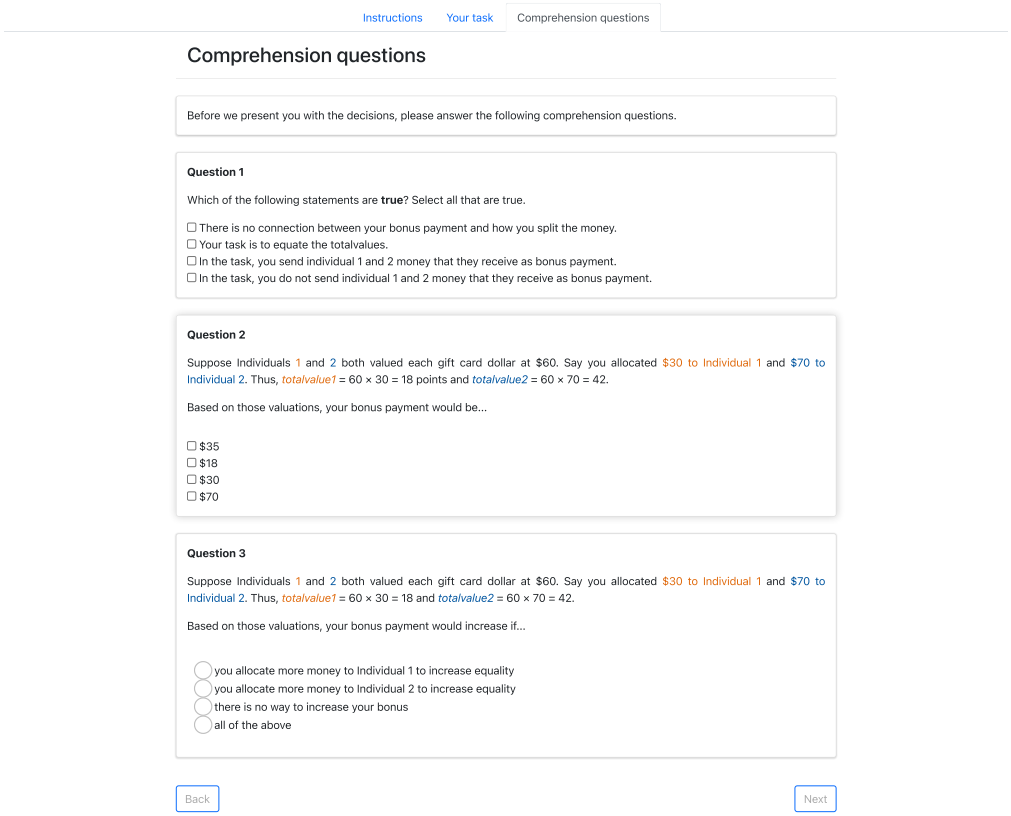


Figure I.41: Ingroup non-social minimum screen 8

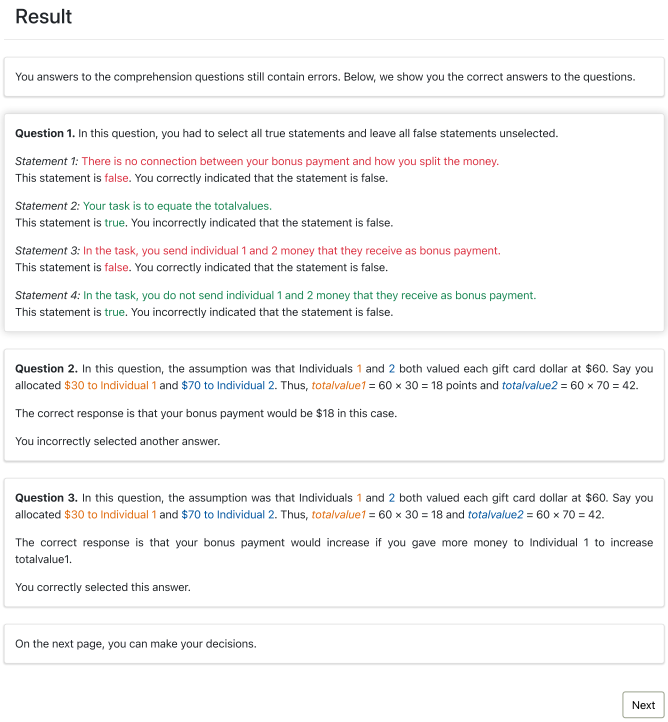


Figure I.42: Ingroup non-social minimum screen 9

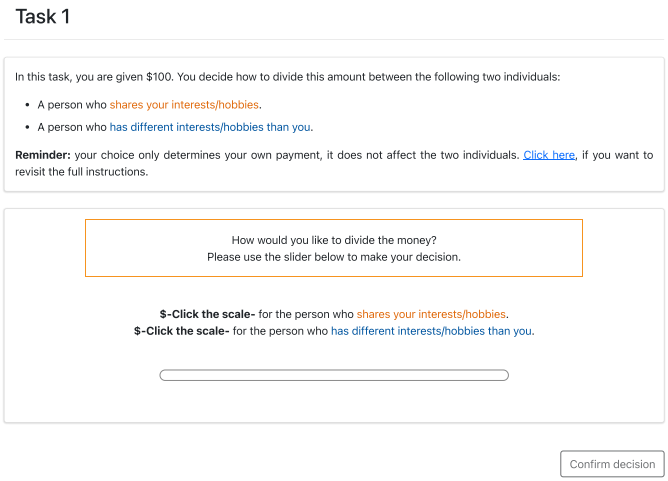


Figure I.43: Ingroup non-social minimum screen 10

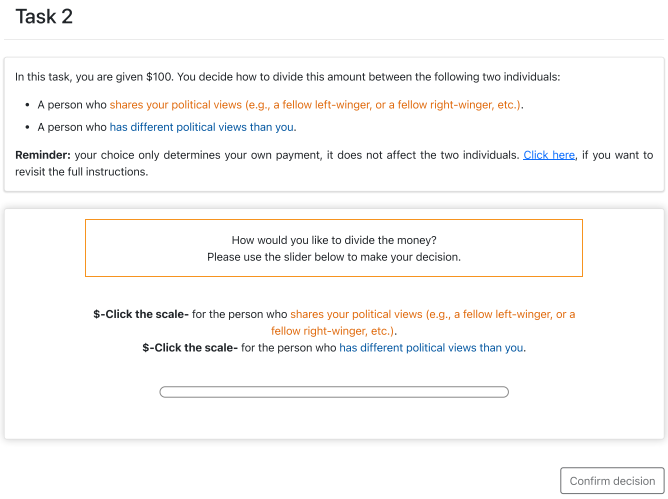
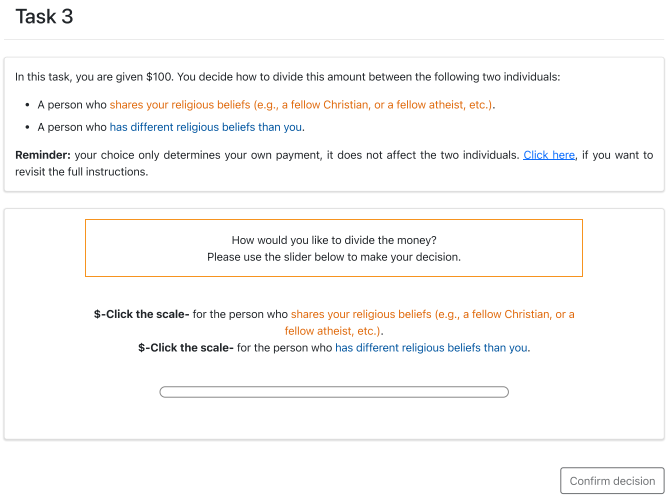


Figure I.44: Ingroup non-social minimum screen 11



I.1.4 Ingroup effort social and Ingroup effort non-social

Figure I.45: Ingroup effort social and Ingroup effort non-social screen 0

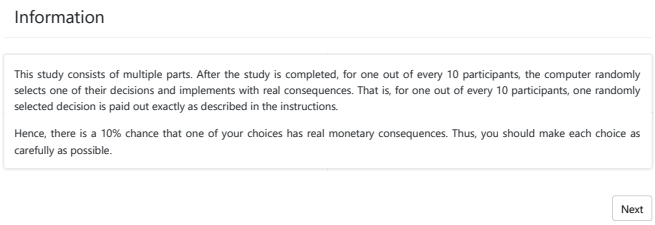


Figure I.46: Ingroup effort social and Ingroup effort non-social screen 2

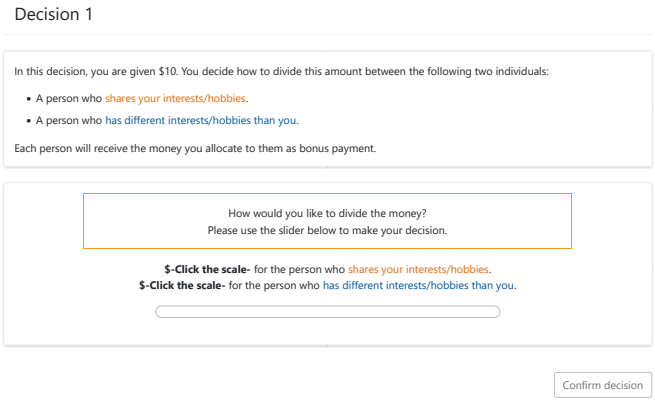


Figure I.47: Ingroup effort social and Ingroup effort non-social screen 4

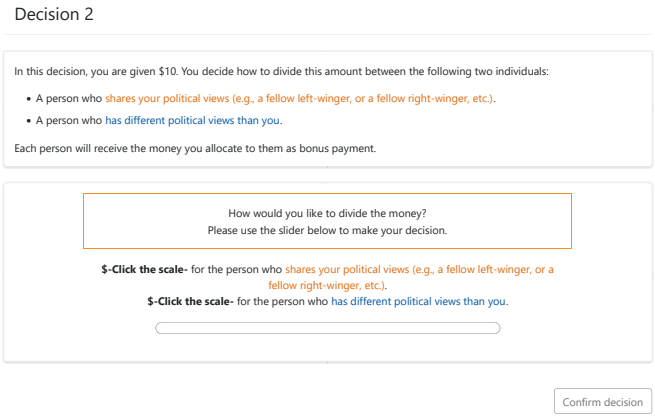


Figure I.48: Ingroup effort social and Ingroup effort non-social screen 6

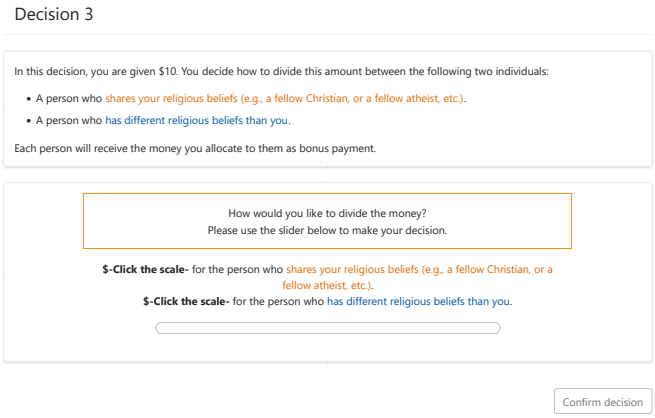


Figure I.49: Ingroup effort social and Ingroup effort non-social screen 8

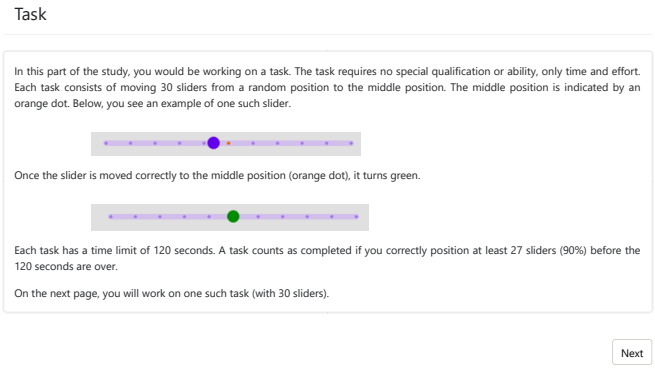


Figure I.50: Ingroup effort social and Ingroup effort non-social screen 10

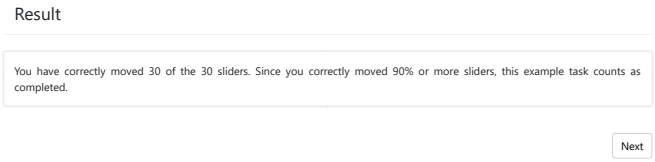


Figure I.51: Ingroup effort social and Ingroup effort non-social screen 12

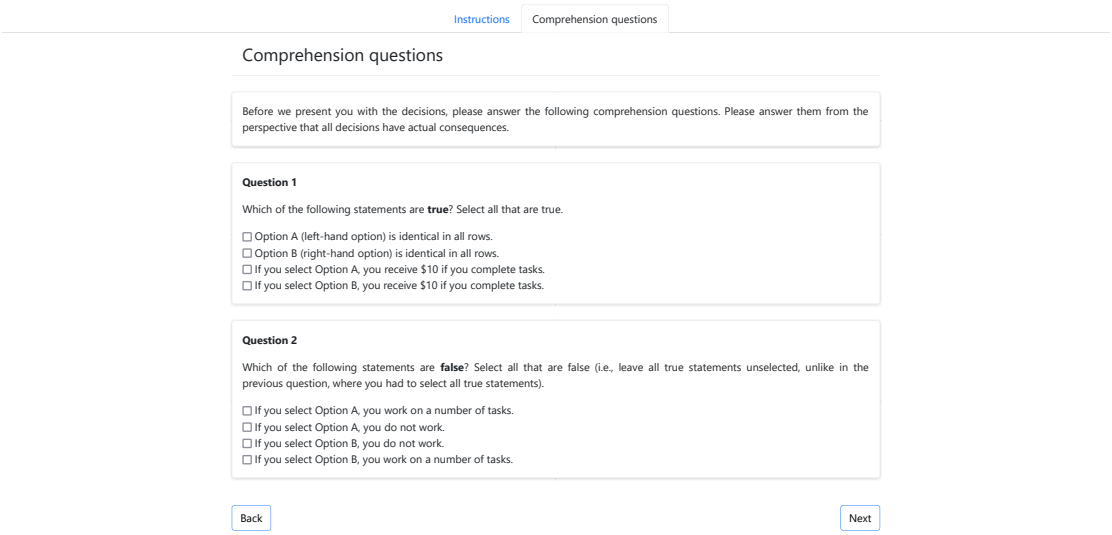


Figure I.52: Ingroup effort social and Ingroup effort non-social screen 14

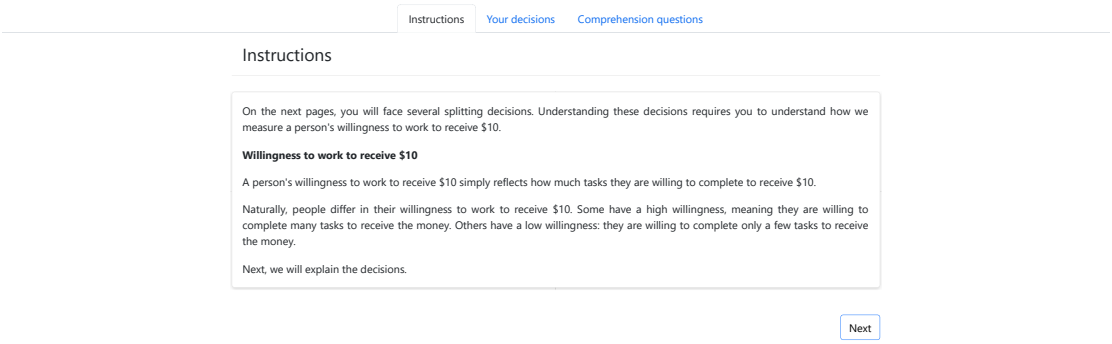


Figure I.53: Ingroup effort social and Ingroup effort non-social screen 16

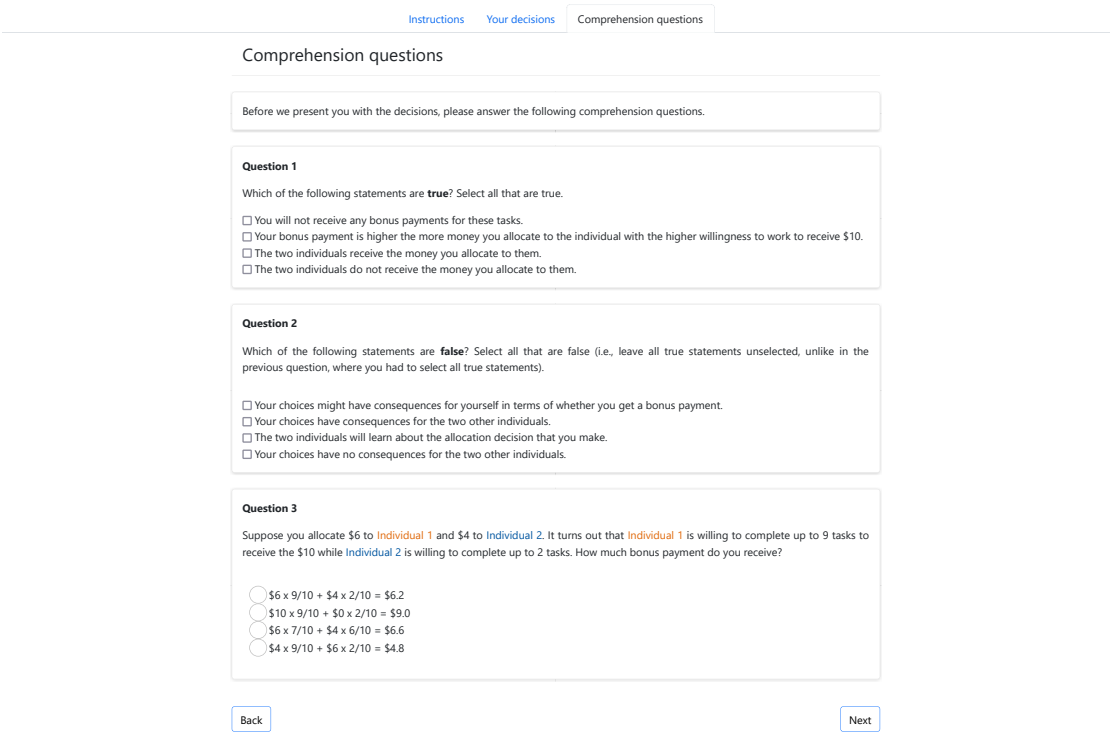
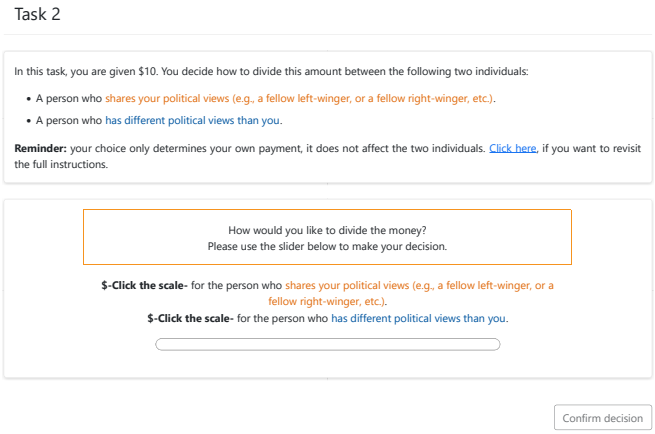


Figure I.54: Ingroup effort social and Ingroup effort non-social screen 18



I.1.5 Ingroup uncertainty

Figure I.55: Ingroup uncertainty screen 1

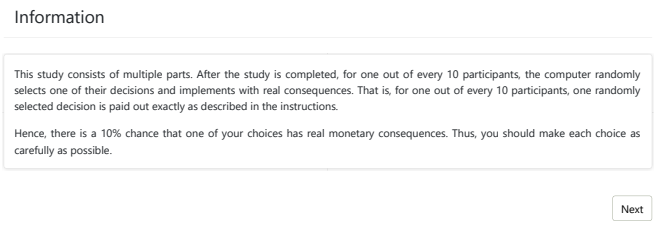


Figure I.56: Ingroup uncertainty screen 2

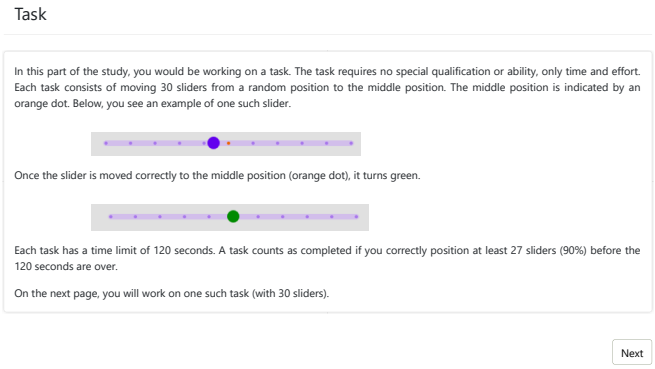


Figure I.57: Ingroup uncertainty screen 3

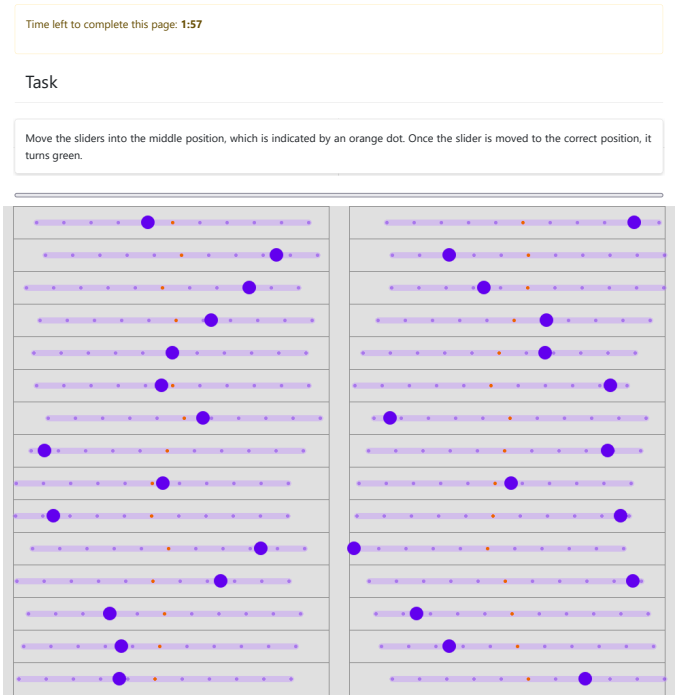


Figure I.58: Ingroup uncertainty screen 4

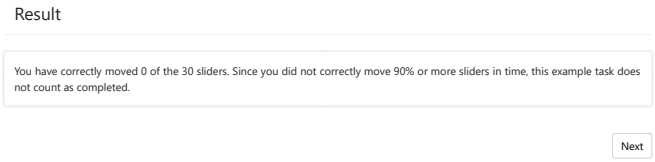


Figure I.59: Ingroup uncertainty screen 5

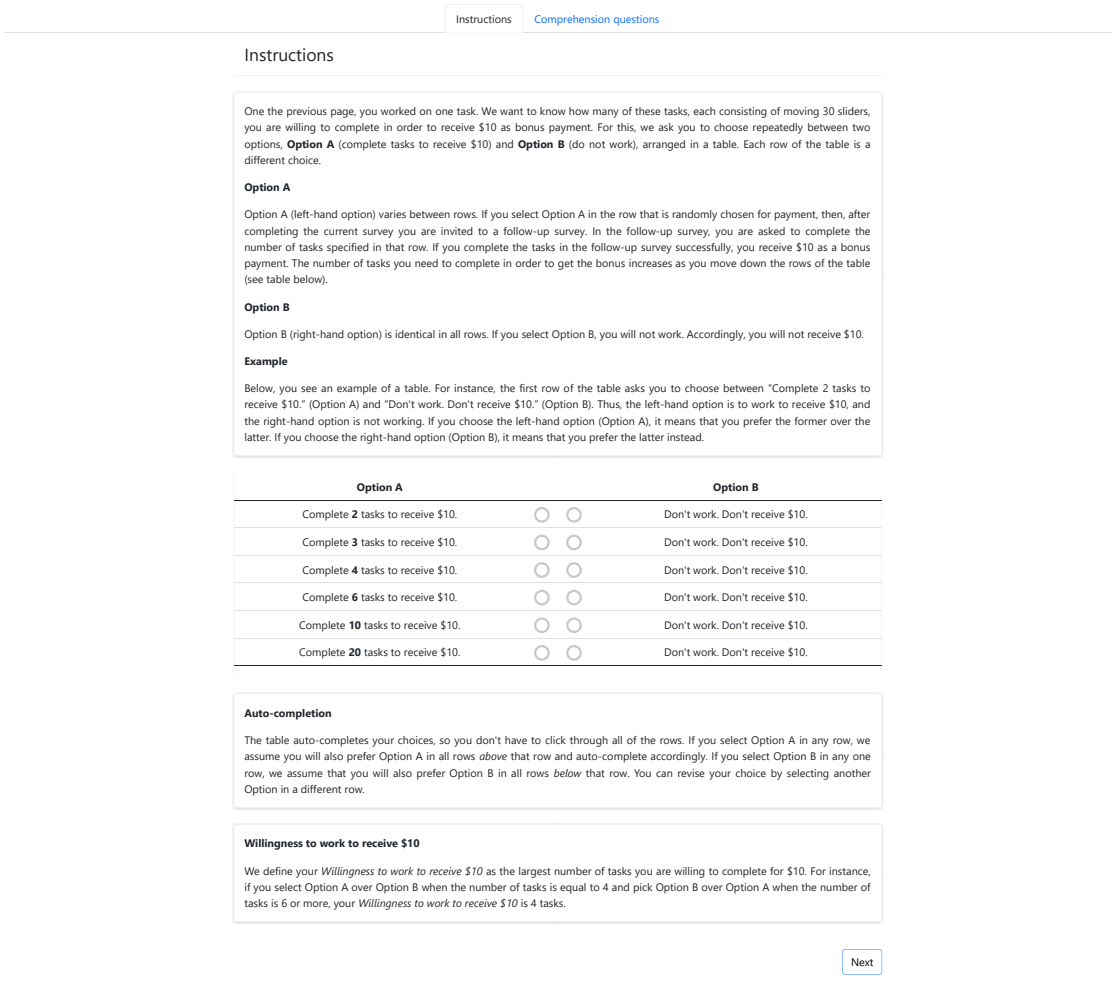


Figure I.60: Ingroup uncertainty screen 6

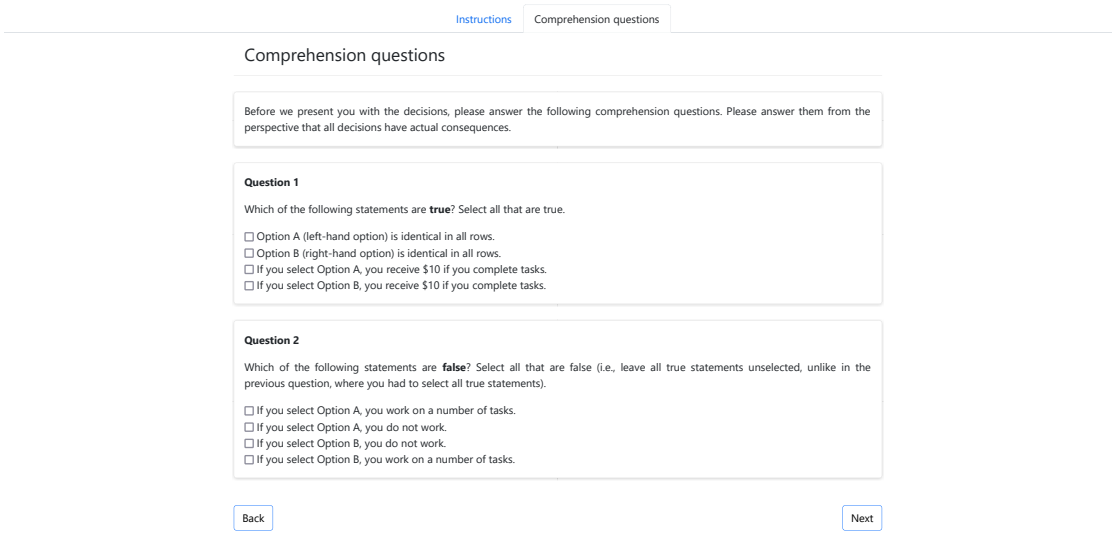


Figure I.61: Ingroup uncertainty screen 7

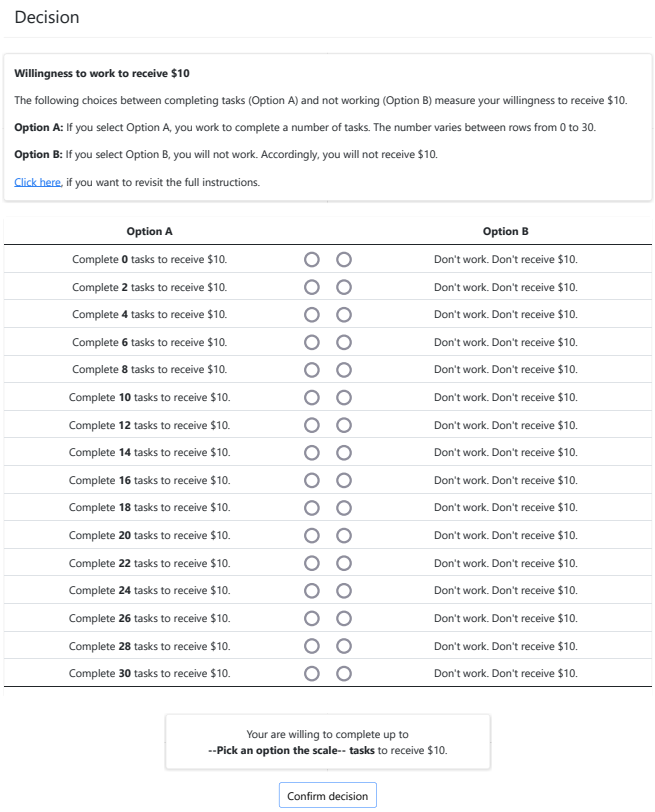


Figure I.62: Ingroup uncertainty screen 8

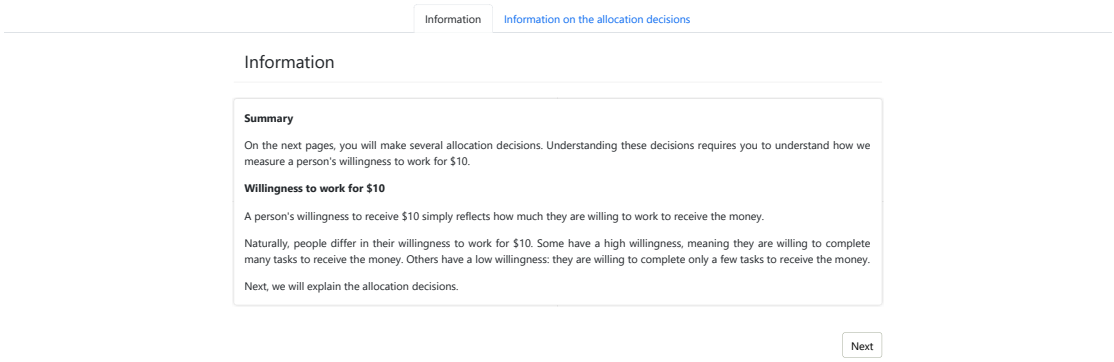


Figure I.63: Ingroup uncertainty screen 9

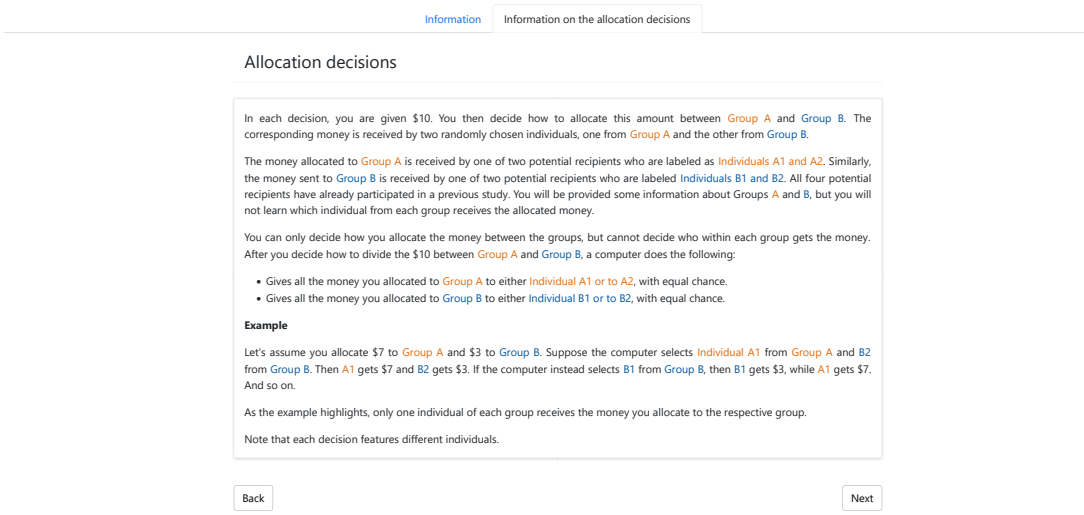


Figure I.64: Ingroup uncertainty screen 10

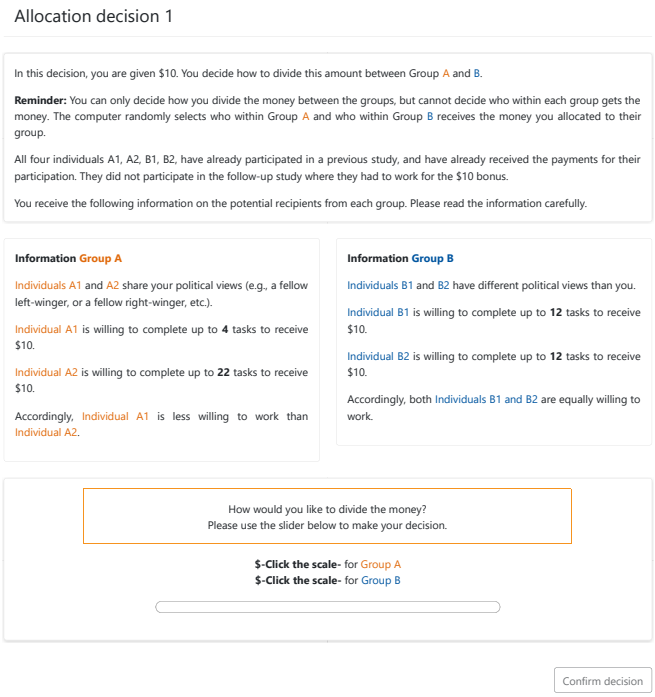


Figure I.65: Ingroup uncertainty screen 11

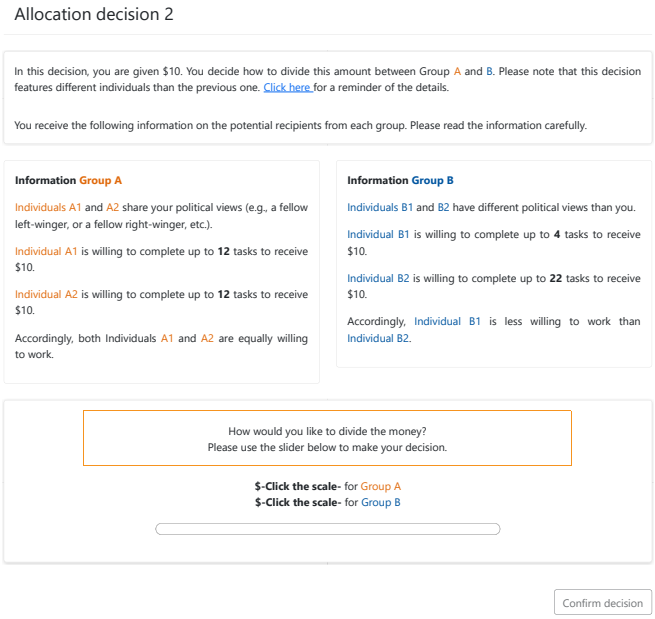


Figure I.66: Ingroup uncertainty screen 12

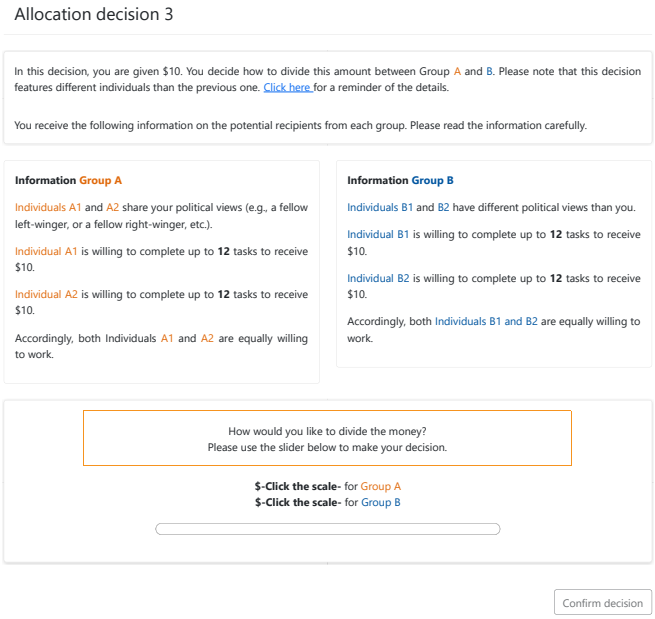


Figure I.67: Ingroup uncertainty screen 13

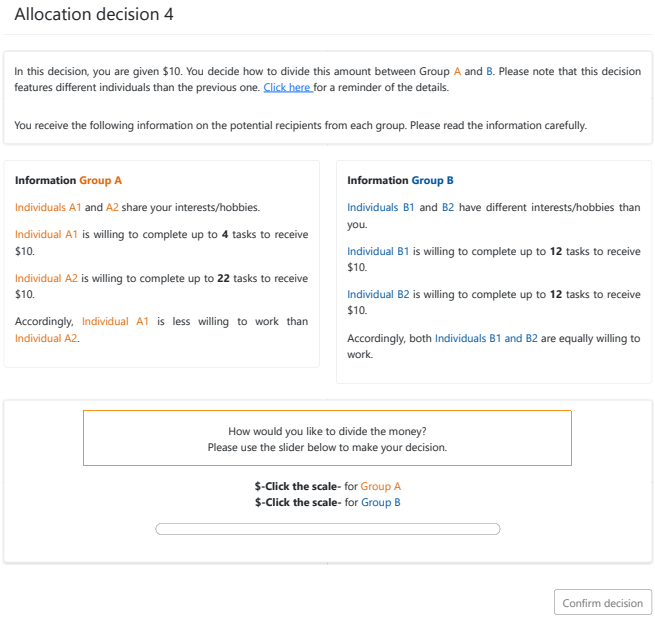


Figure I.68: Ingroup uncertainty screen 14

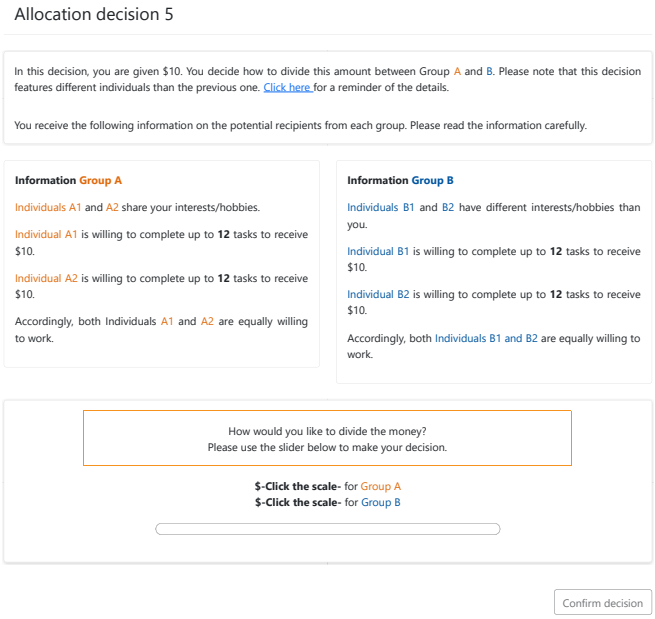


Figure I.69: Ingroup uncertainty screen 15

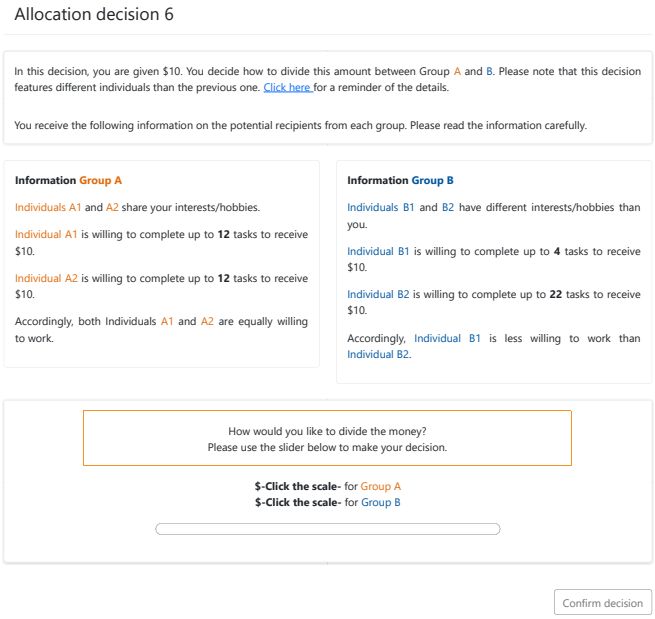


Figure I.70: Ingroup uncertainty screen 16

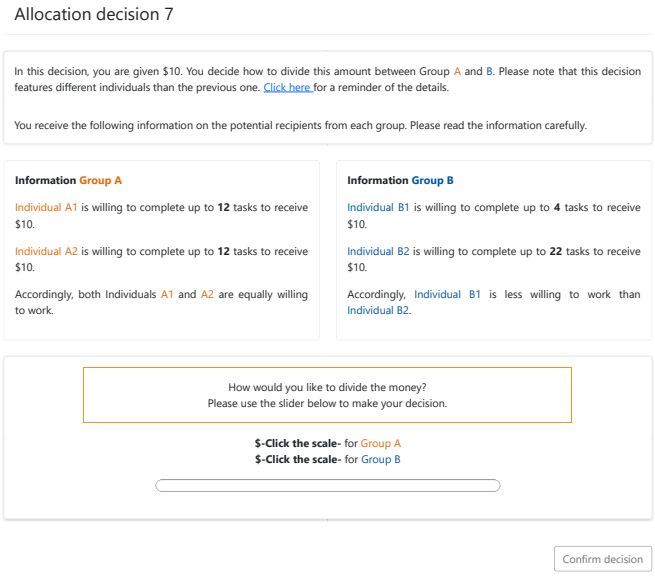


Figure I.71: Ingroup uncertainty screen 17

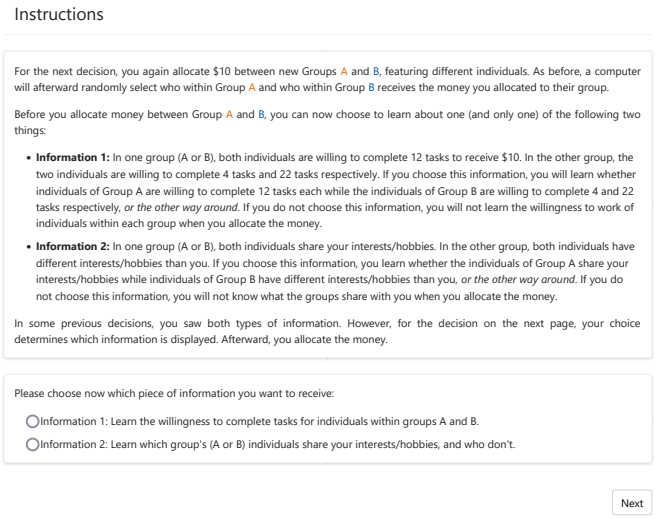
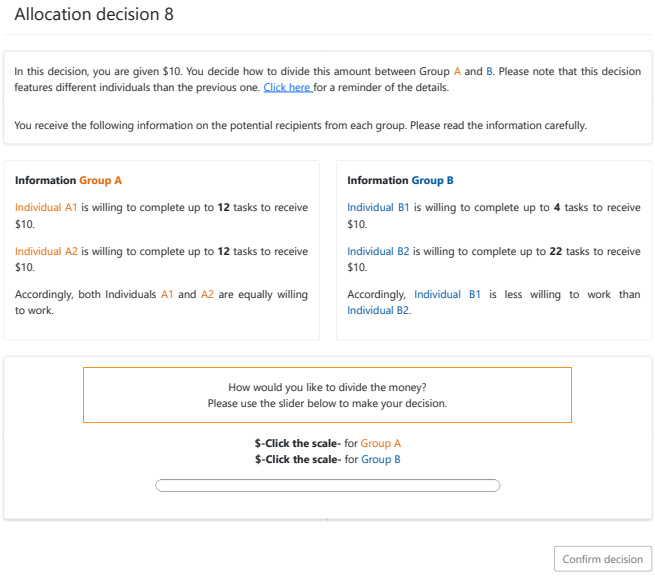


Figure I.72: Ingroup uncertainty screen 18



I.2 Self versus other paradigm screens

I.2.1 Self social and non-social screens

Figure I.73: Self social and non-social screen 1

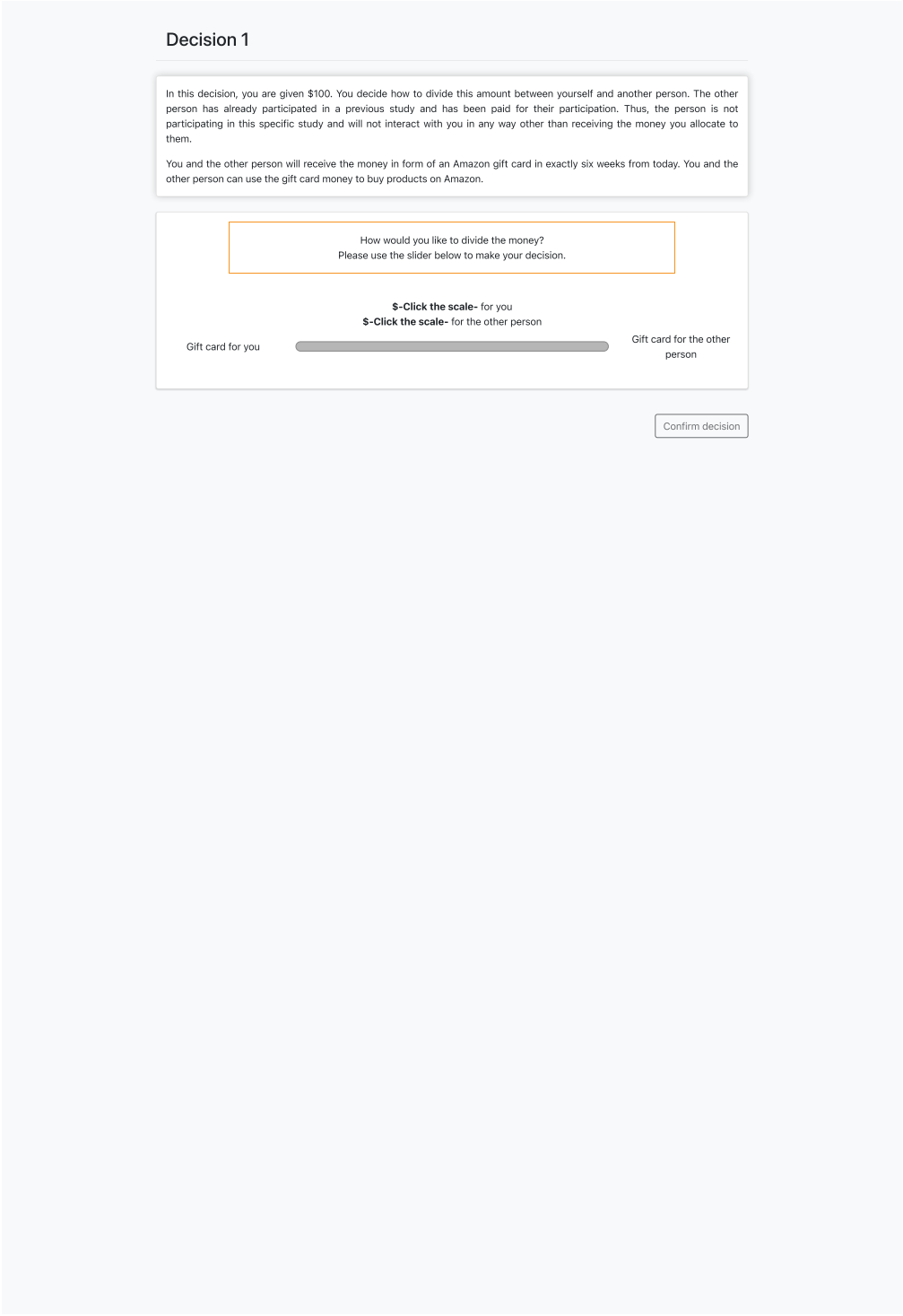


Figure I.74: Self social and non-social screen 2

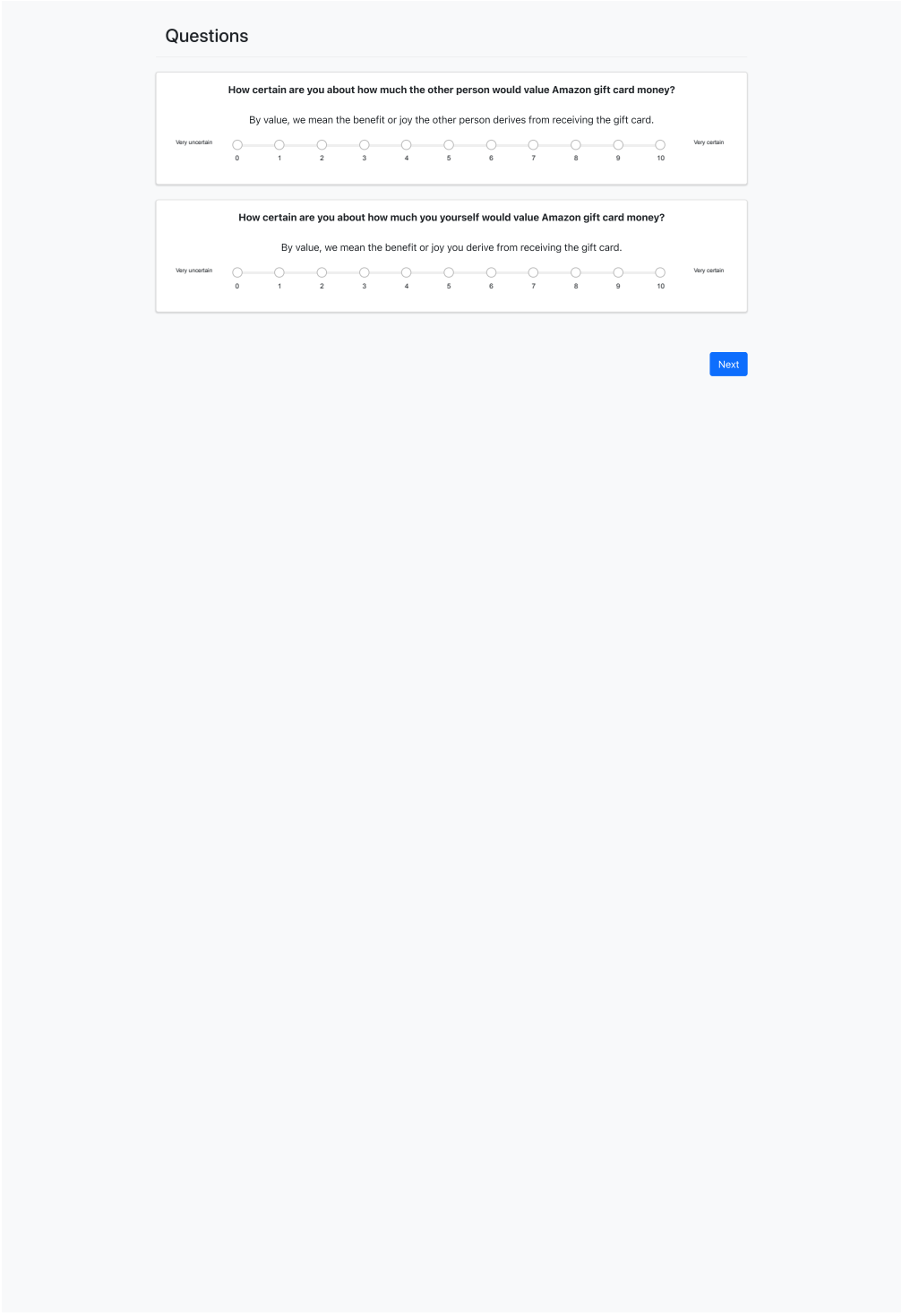


Figure I.75: Self social and non-social screen 3

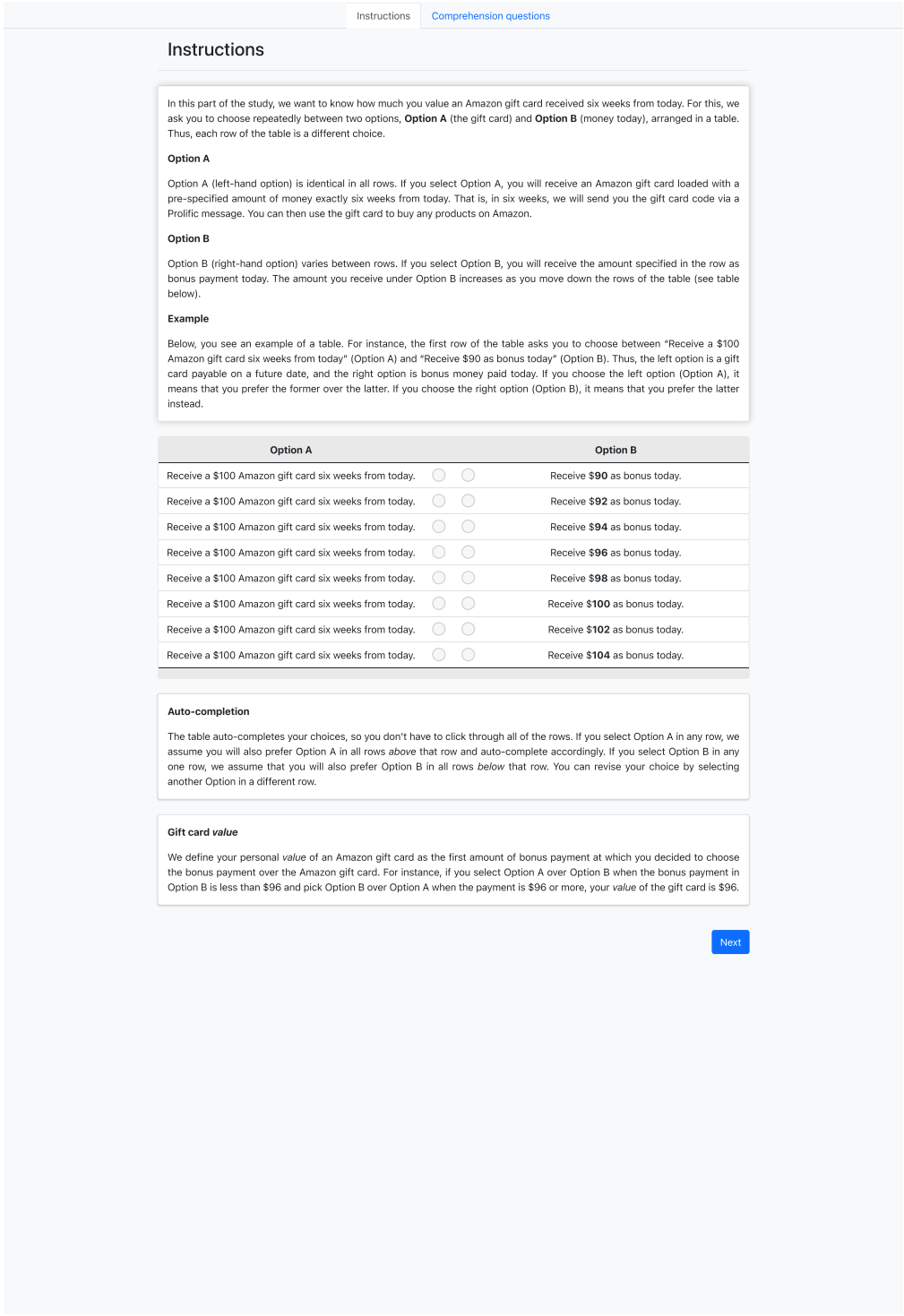


Figure I.76: Self social and non-social screen 4

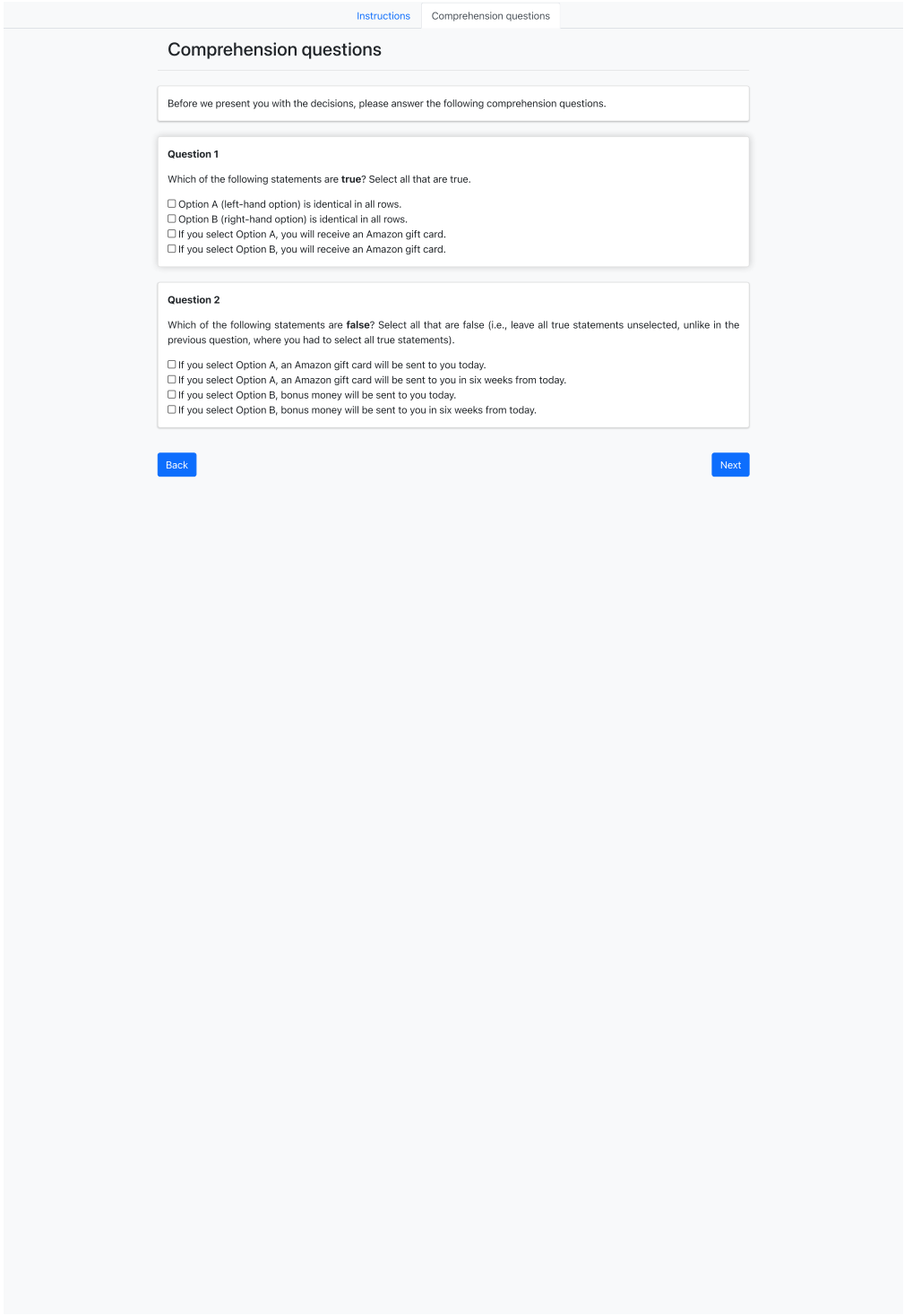


Figure I.77: Self social and non-social screen 5

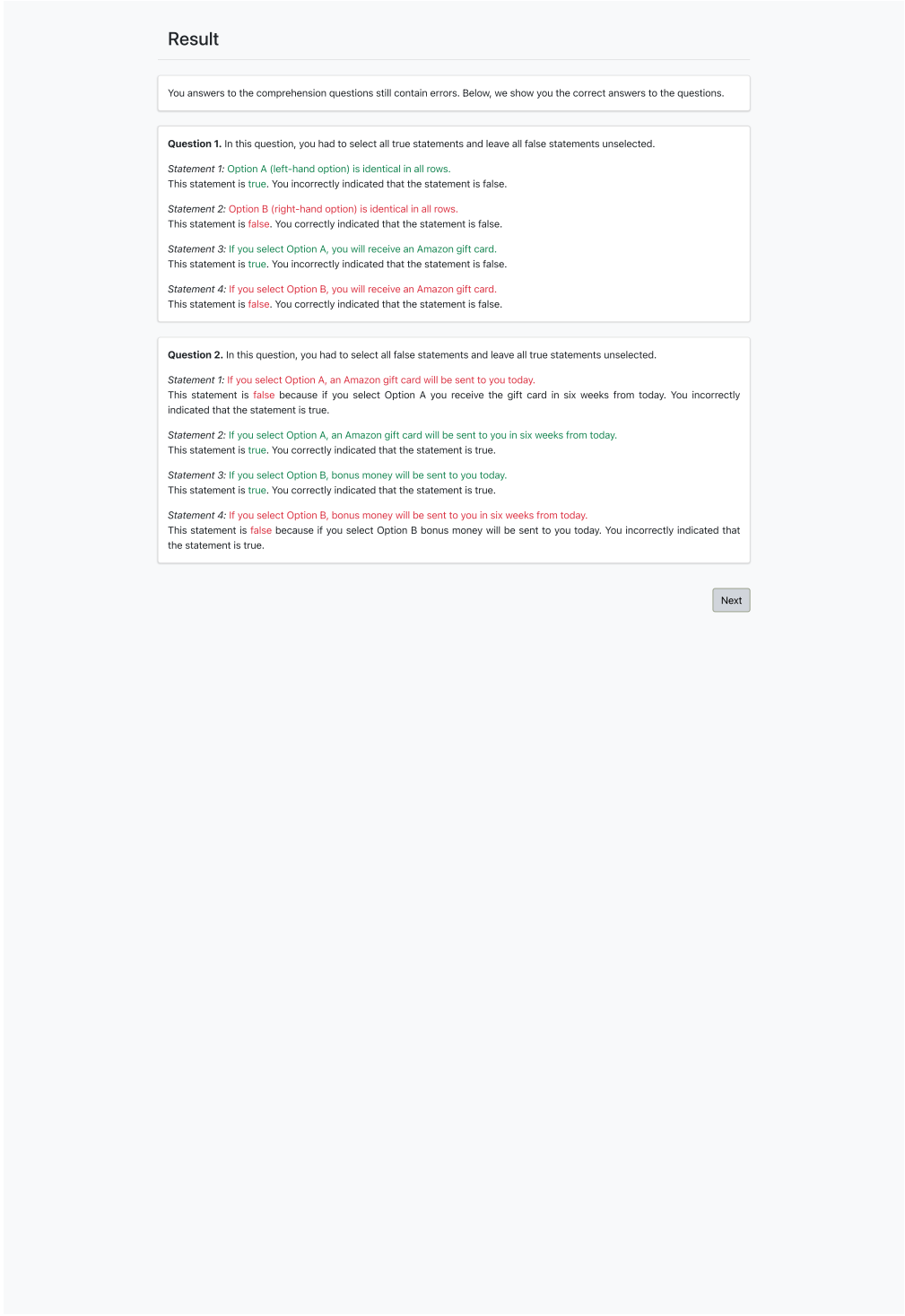


Figure I.78: Self social and non-social screen 6

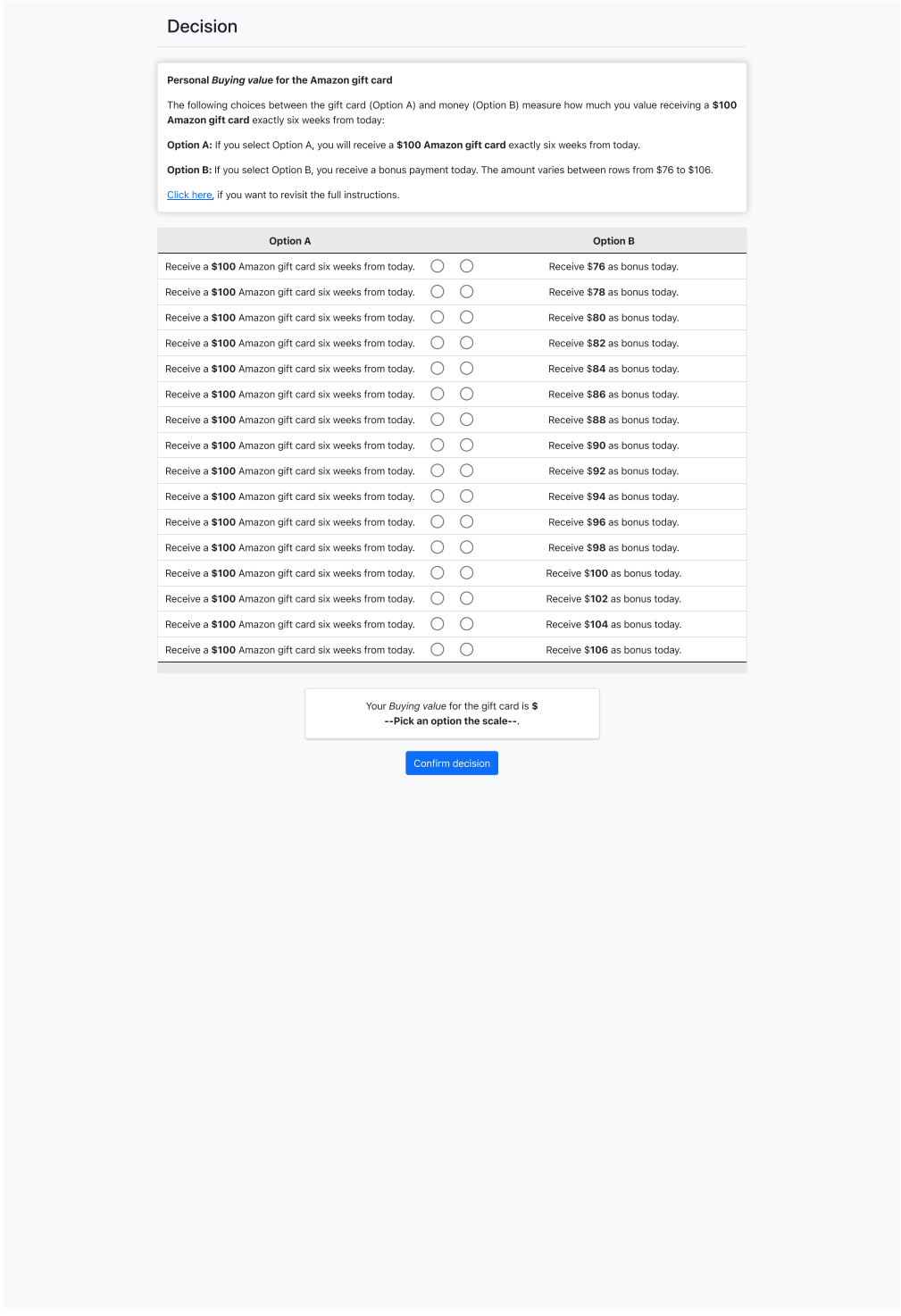


Figure I.79: Self social and non-social screen 7

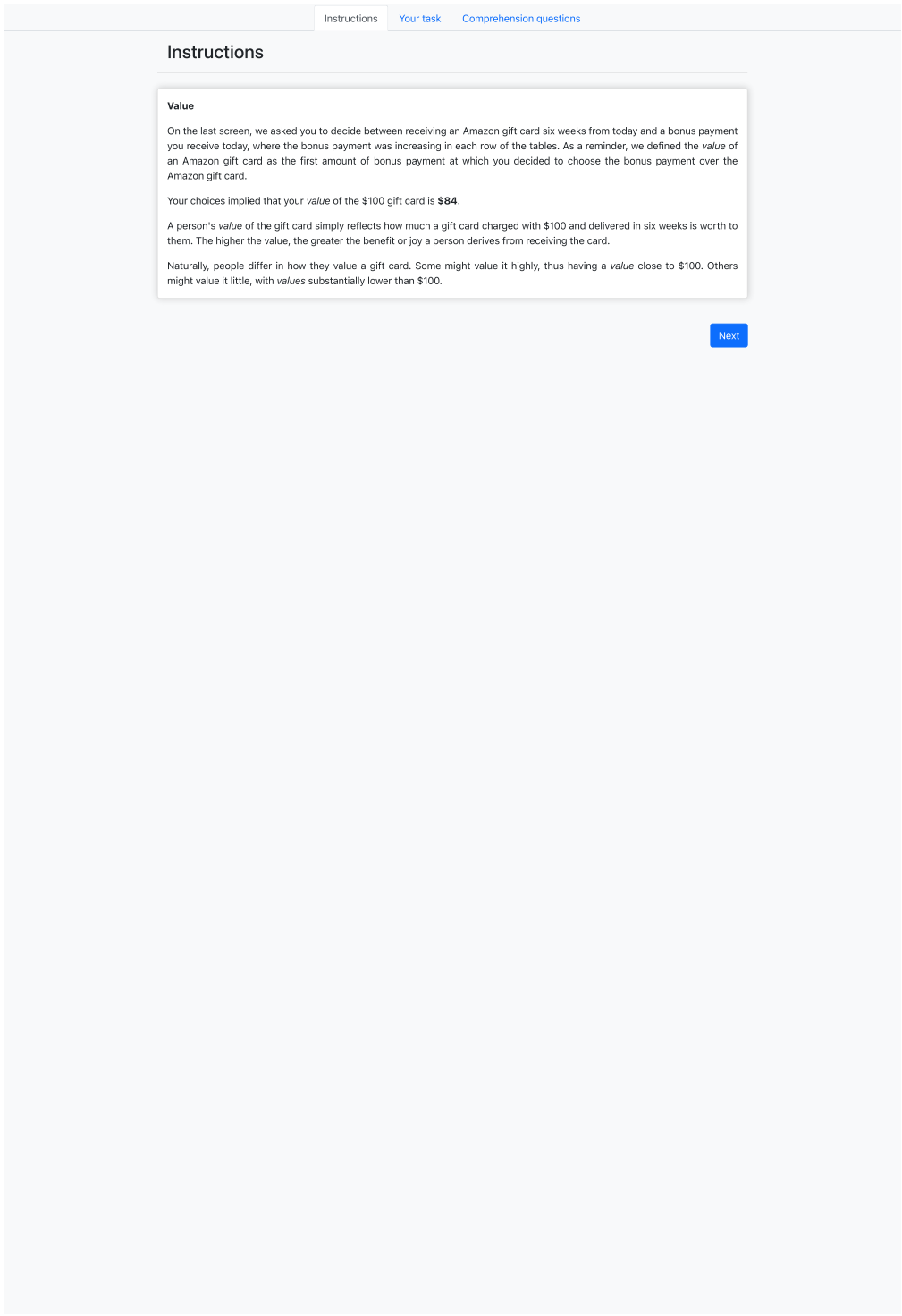


Figure I.80: Self social and non-social screen 8

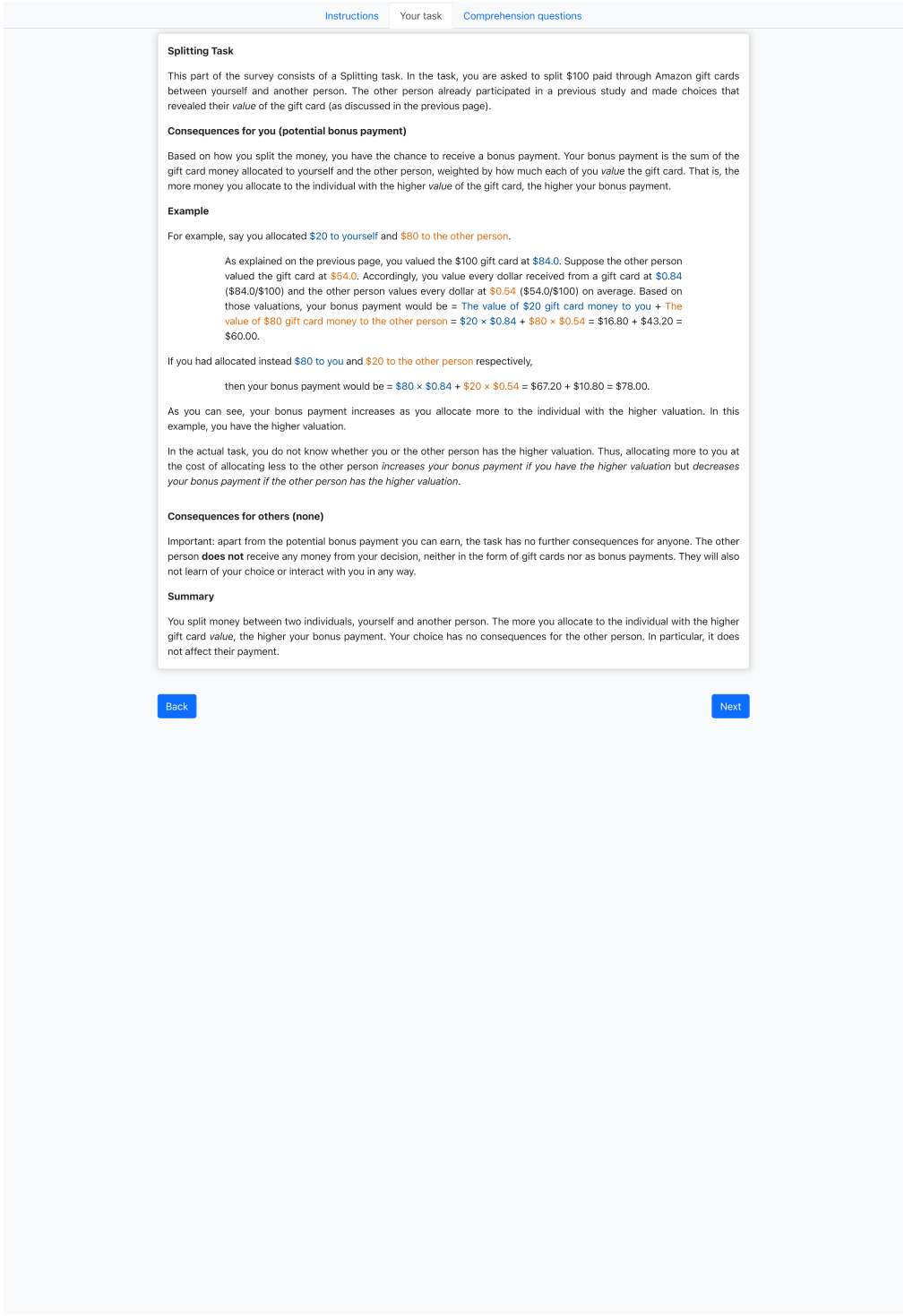


Figure I.81: Self social and non-social screen 9

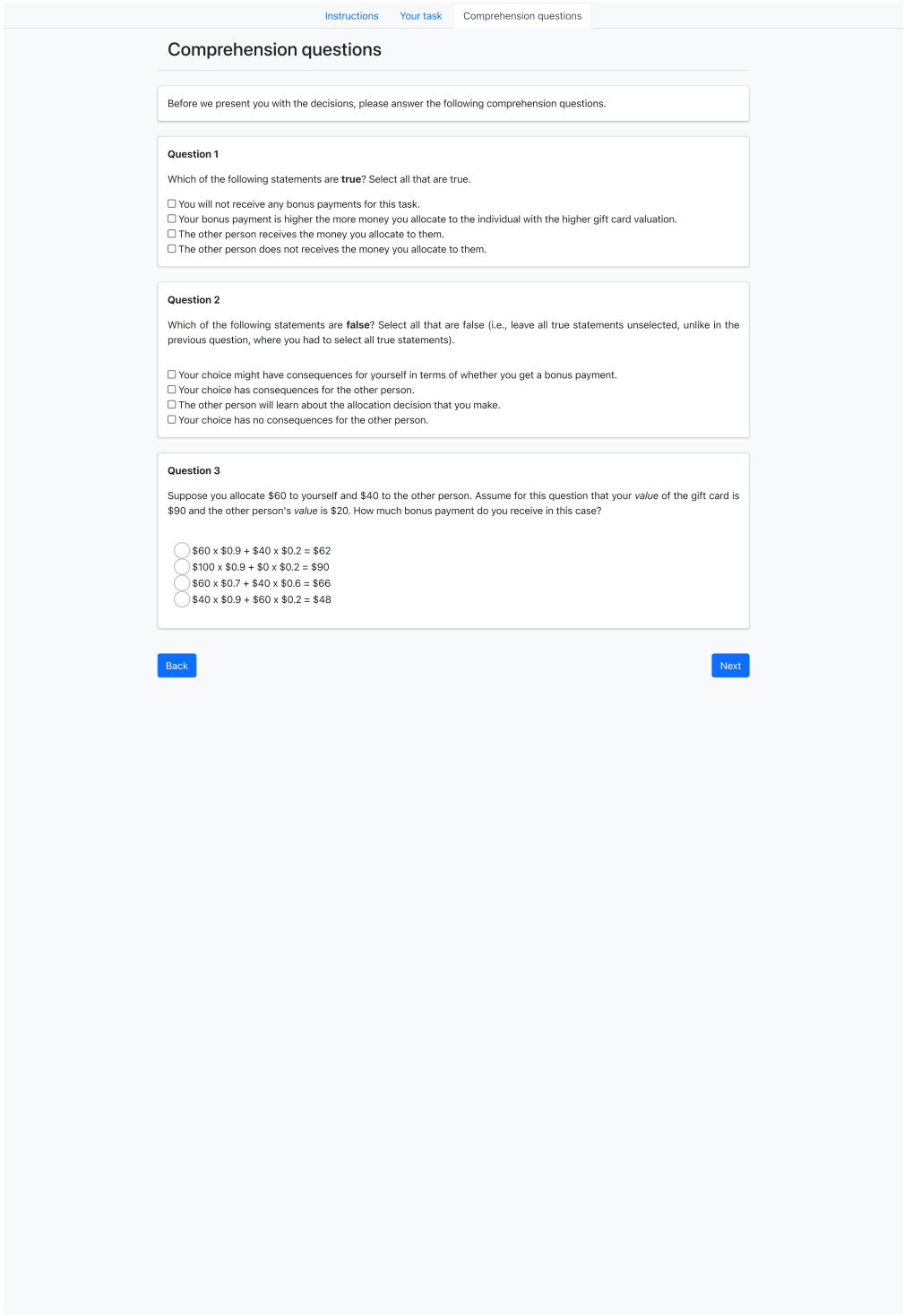


Figure I.82: Self social and non-social screen 10

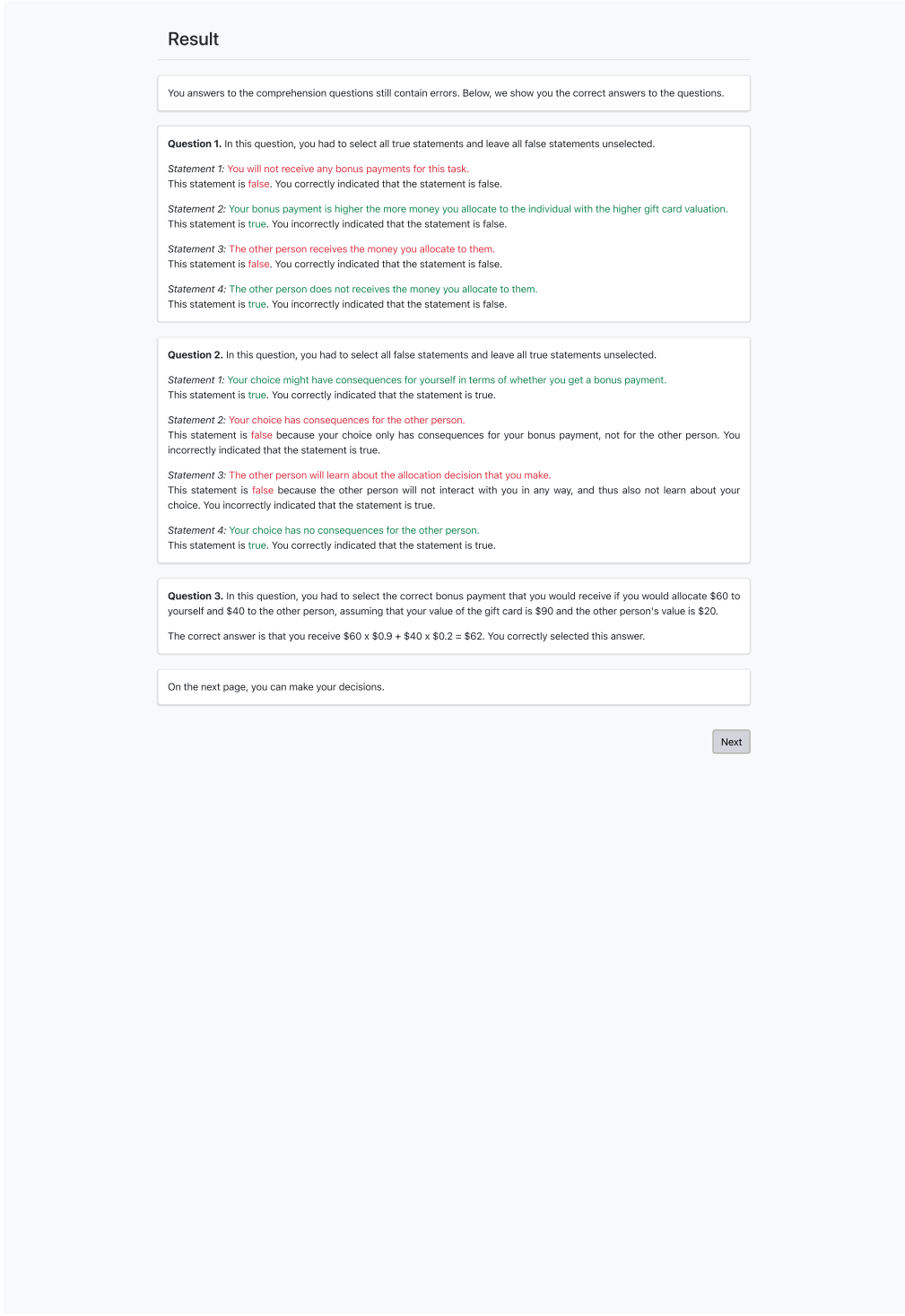


Figure I.83: Self social and non-social screen 11

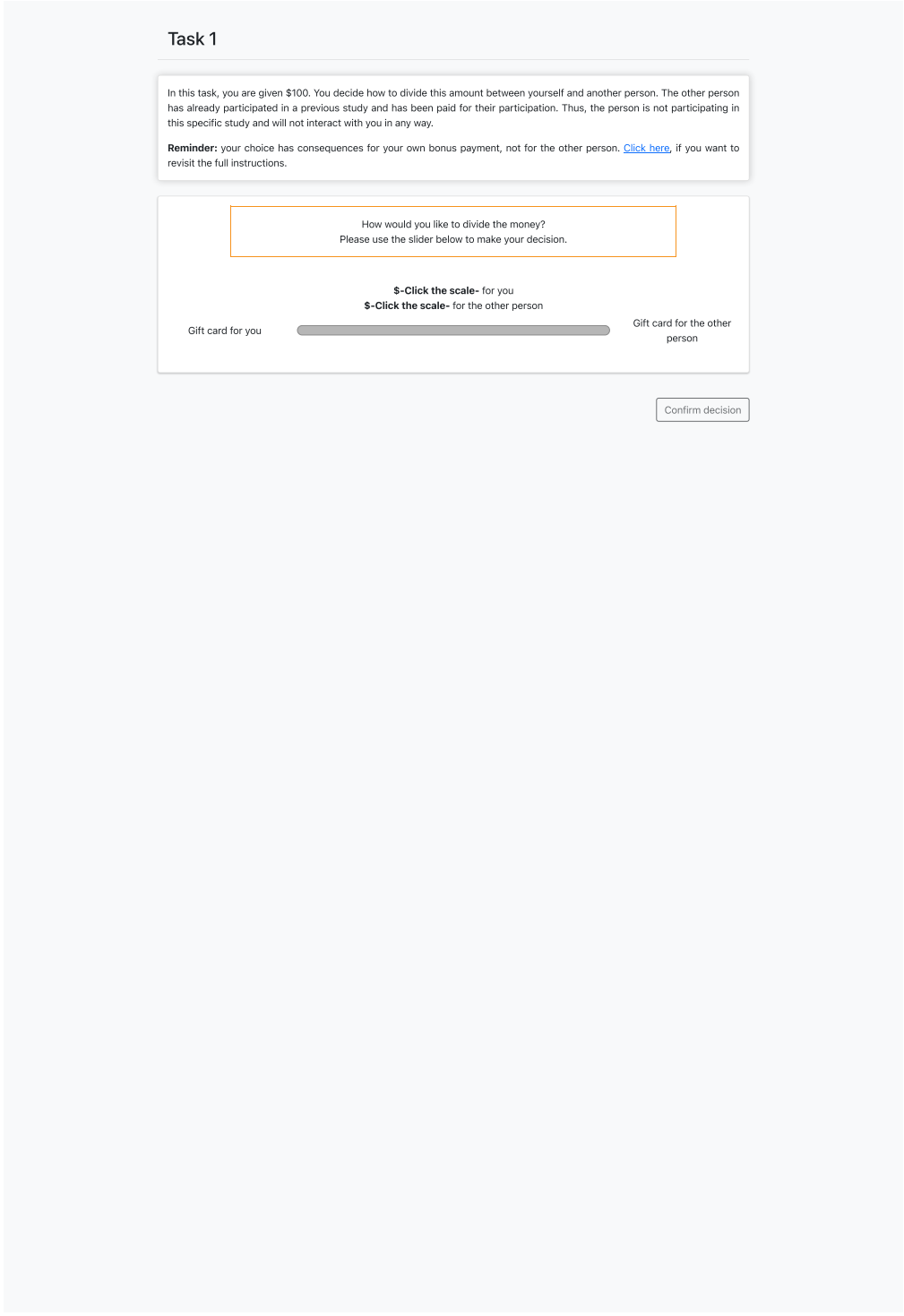
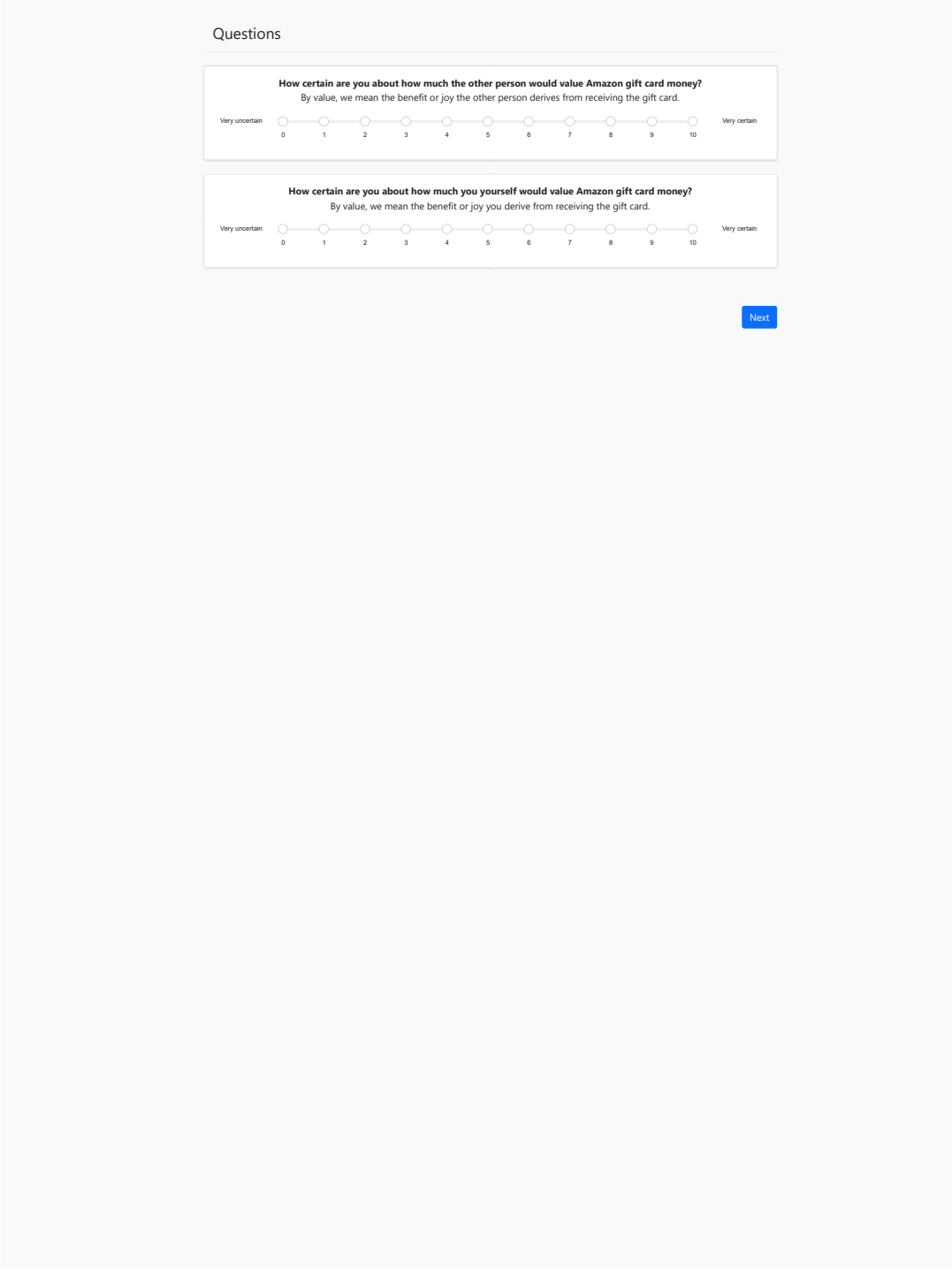


Figure I.84: Self social and non-social screen 12



I.2.2 Incentive self and Incentive other screens

Figure I.85: Incentive self and Incentive other screen 1

Figure I.86: Incentive self and Incentive other screen 2

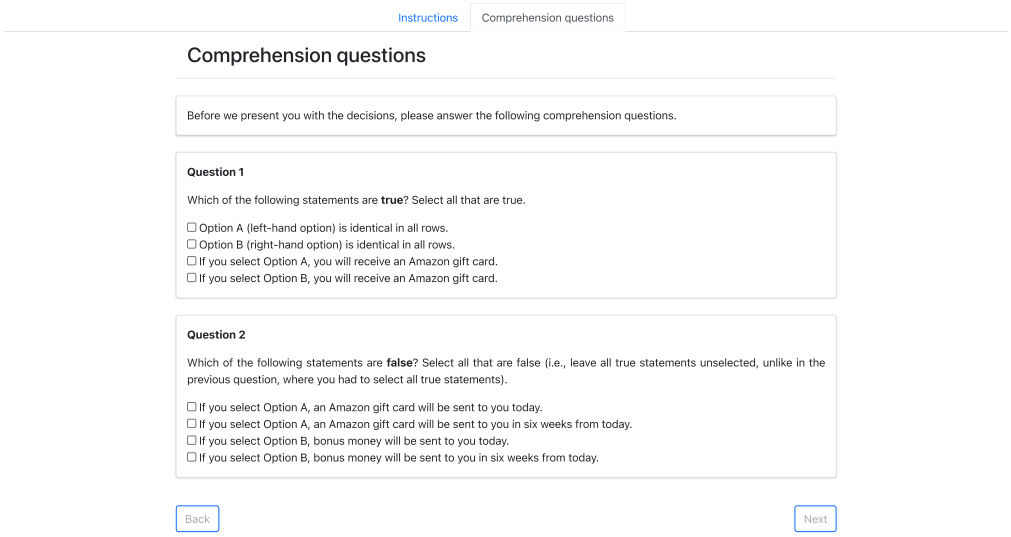


Figure I.87: Incentive self and Incentive other screen 3

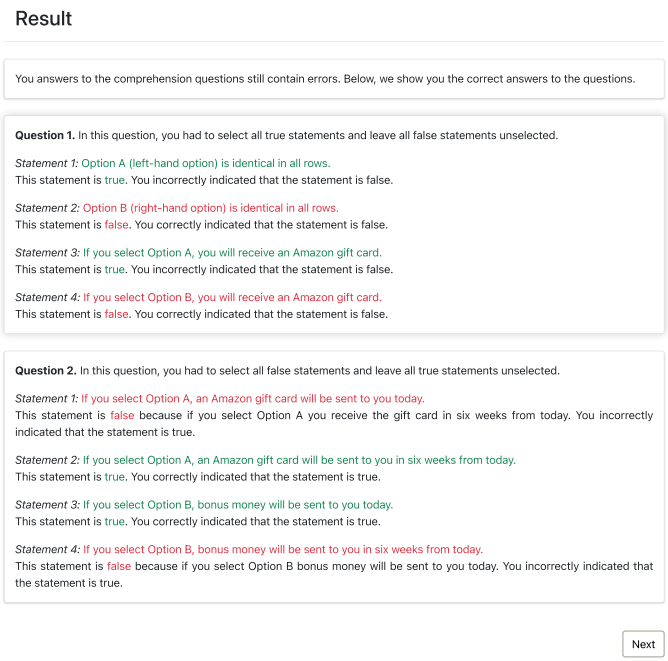


Figure I.88: Incentive self and Incentive other screen 4

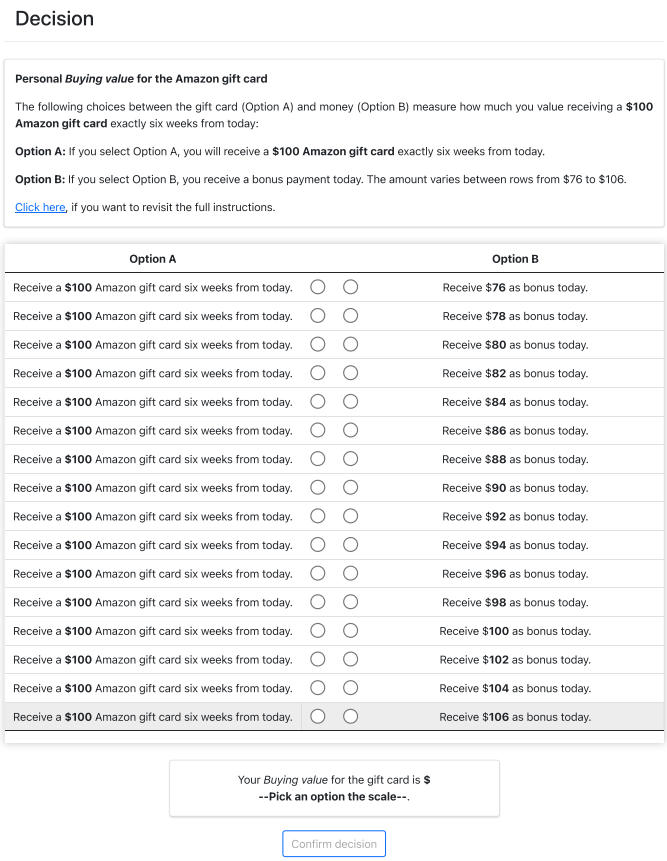


Figure I.89: Incentive self and Incentive other screen 5

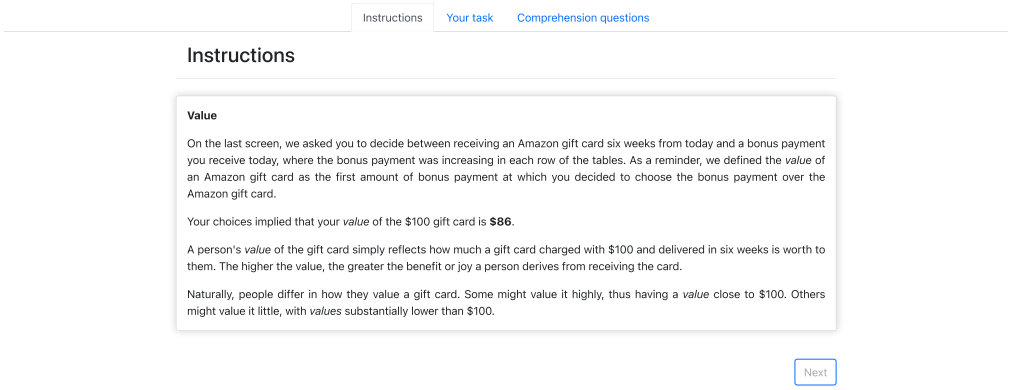


Figure I.90: Incentive self and Incentive other screen 6

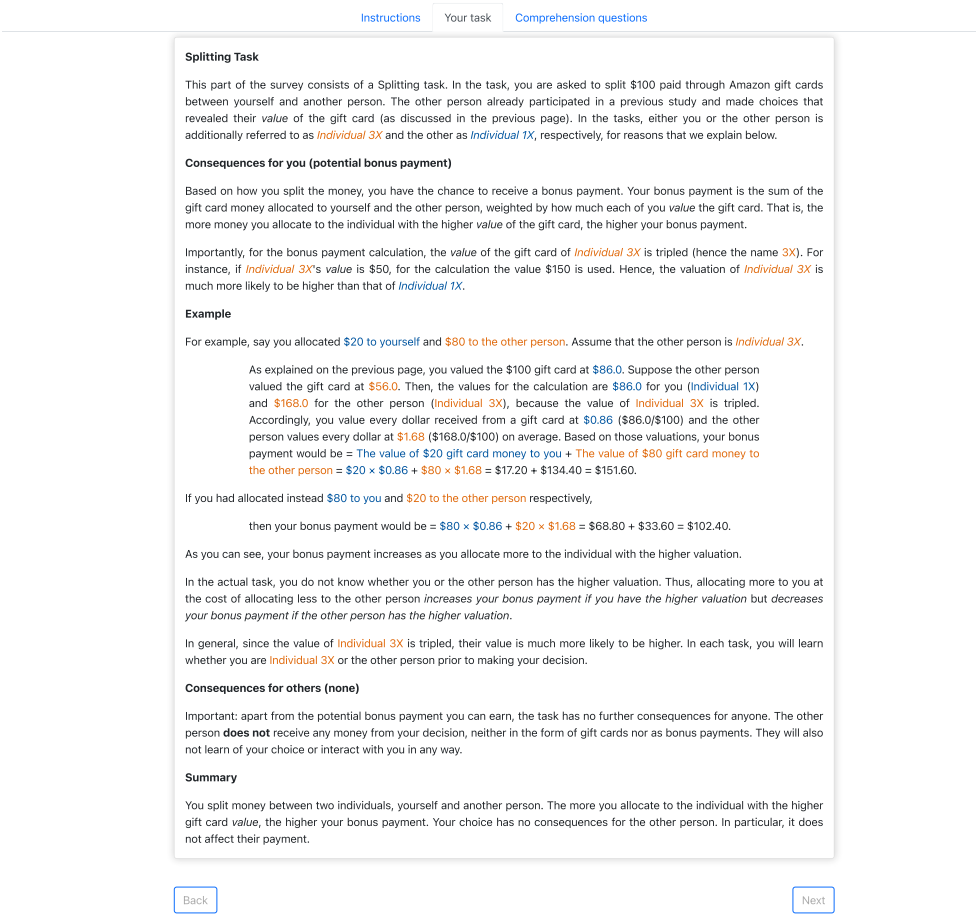


Figure I.91: Incentive self and Incentive other screen 7

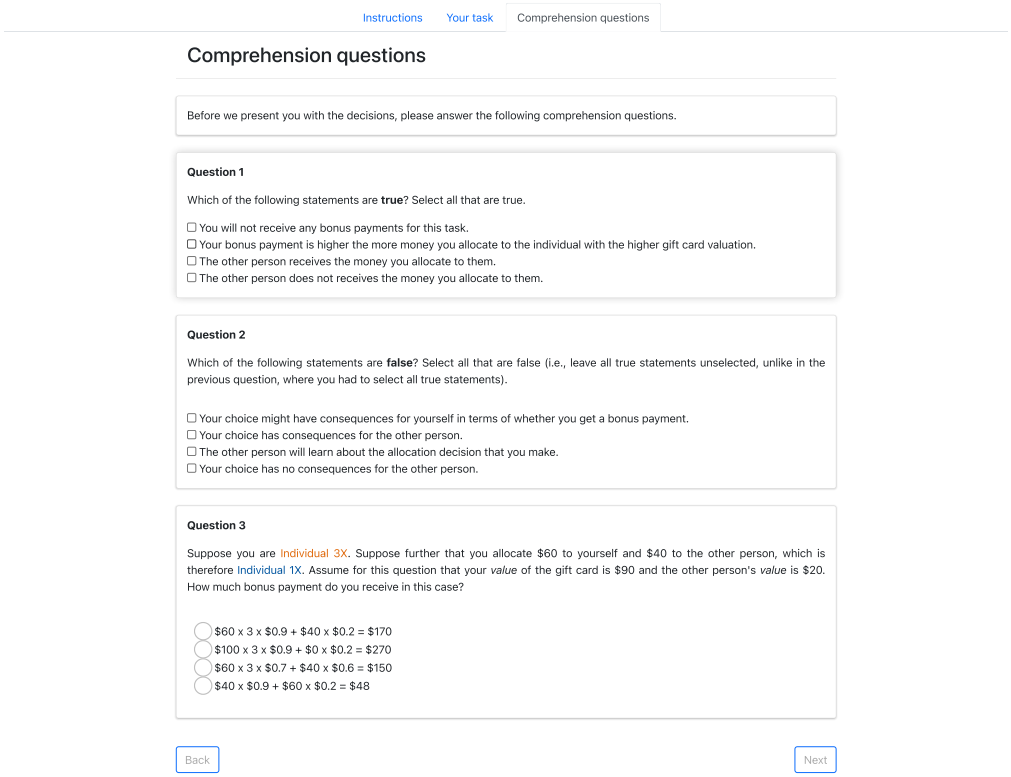


Figure I.92: Incentive self and Incentive other screen 8

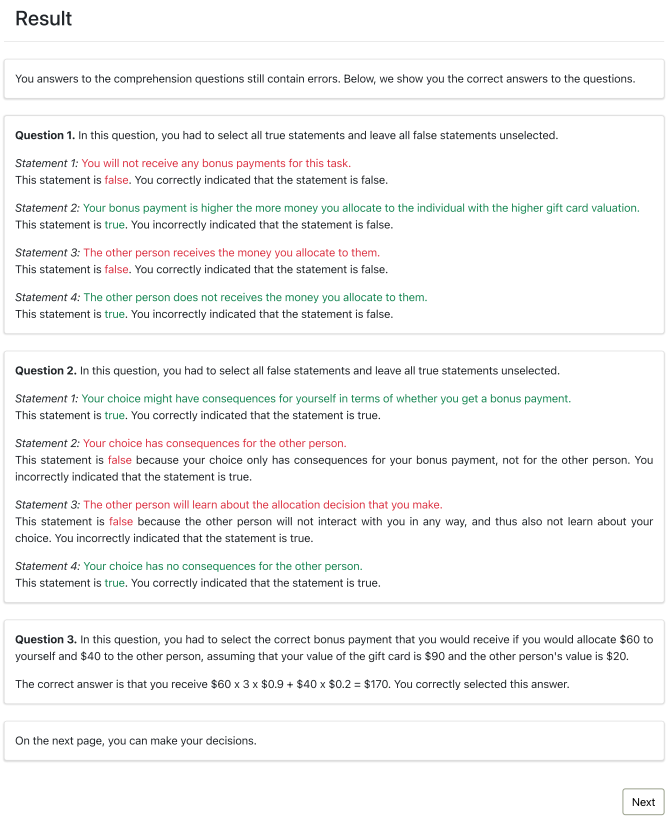


Figure I.93: Incentive self and Incentive other screen 9

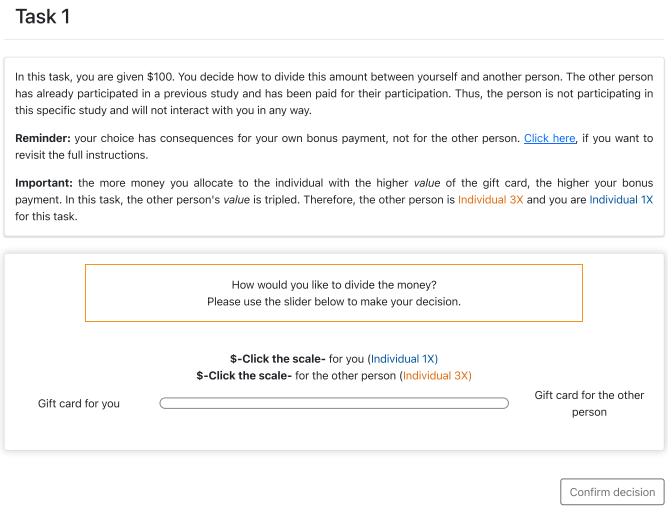


Figure I.94: Incentive self and Incentive other screen 10

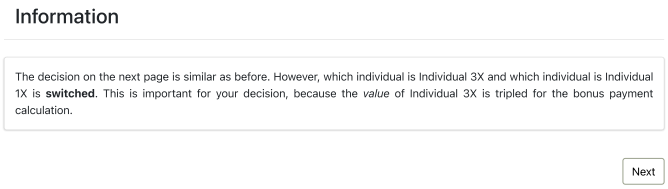
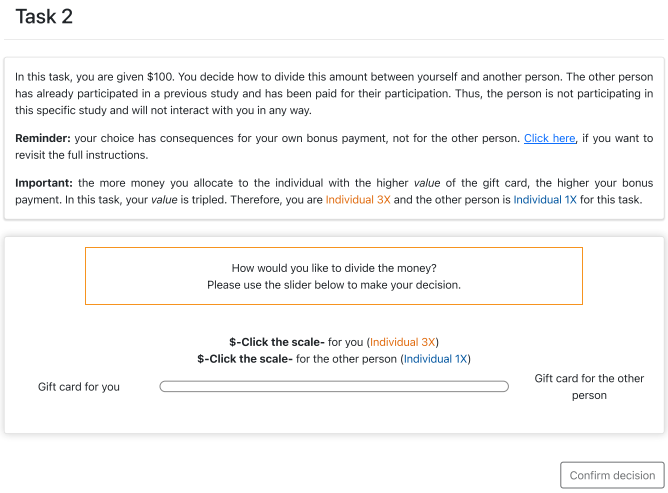


Figure I.95: Incentive self and Incentive other screen 11



I.3 Giving versus Taking paradigm screens

Figure I.96: Giving versus Taking paradigm screen 1

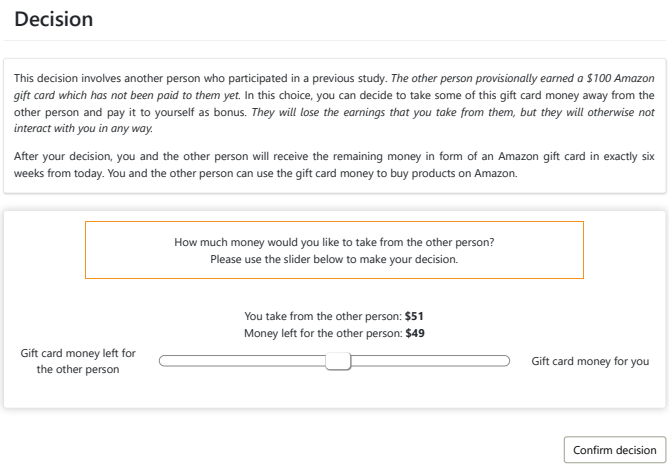


Figure I.97: Giving versus Taking paradigm screen 2

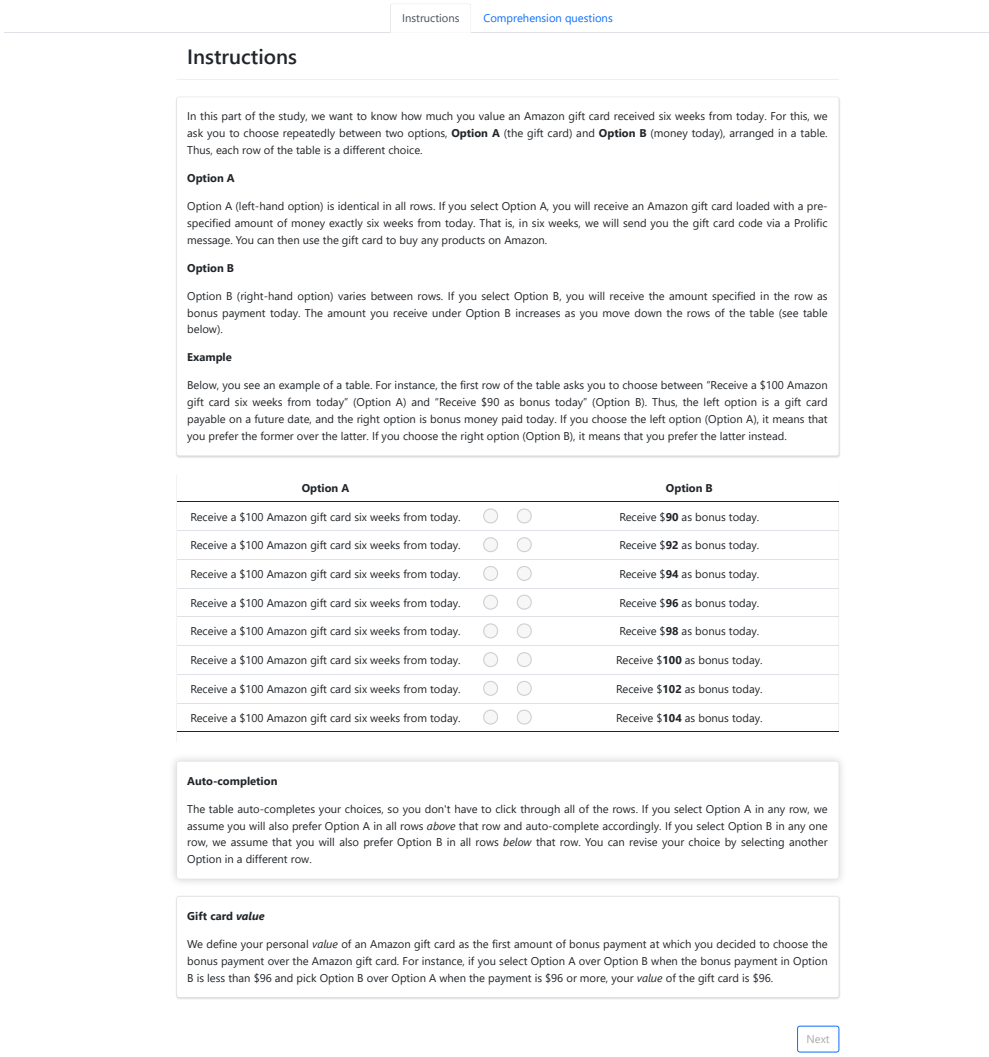


Figure I.98: Giving versus Taking paradigm screen 3

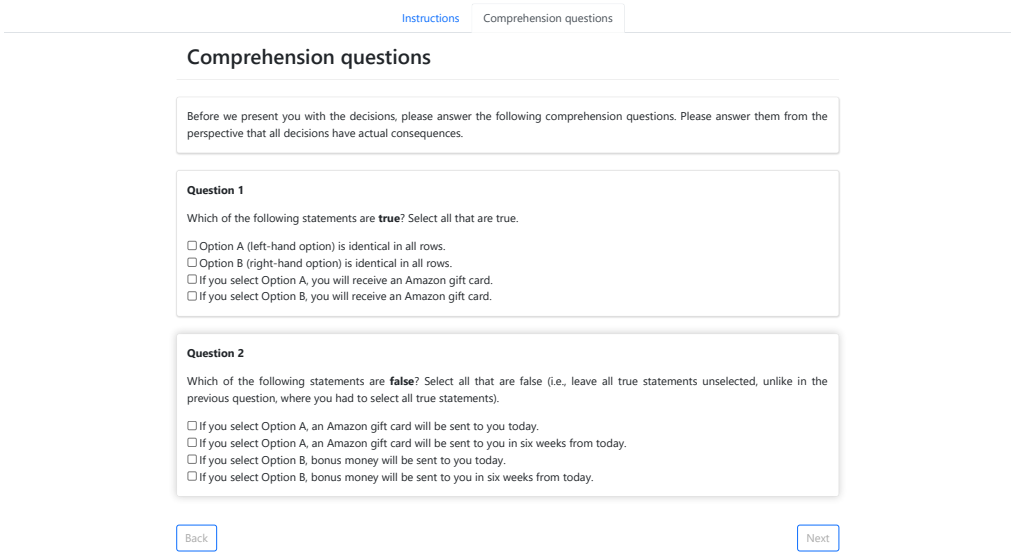


Figure I.99: Giving versus Taking paradigm screen 4

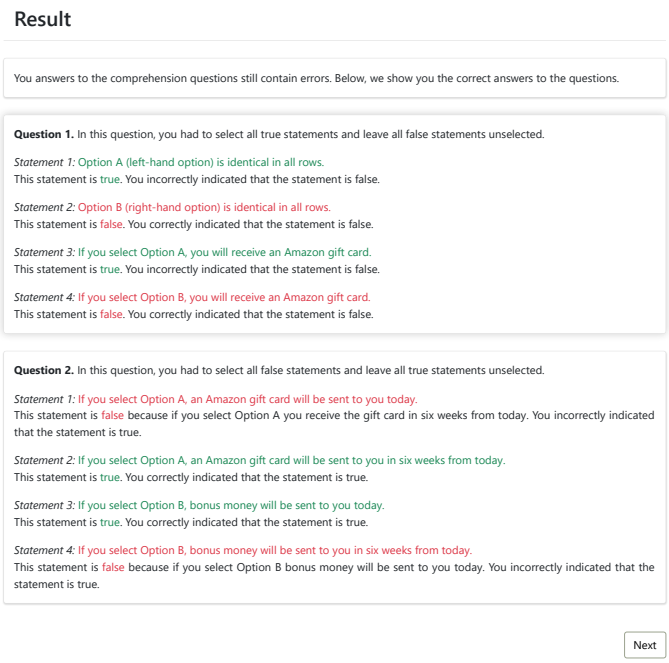


Figure I.100: Giving versus Taking paradigm screen 5

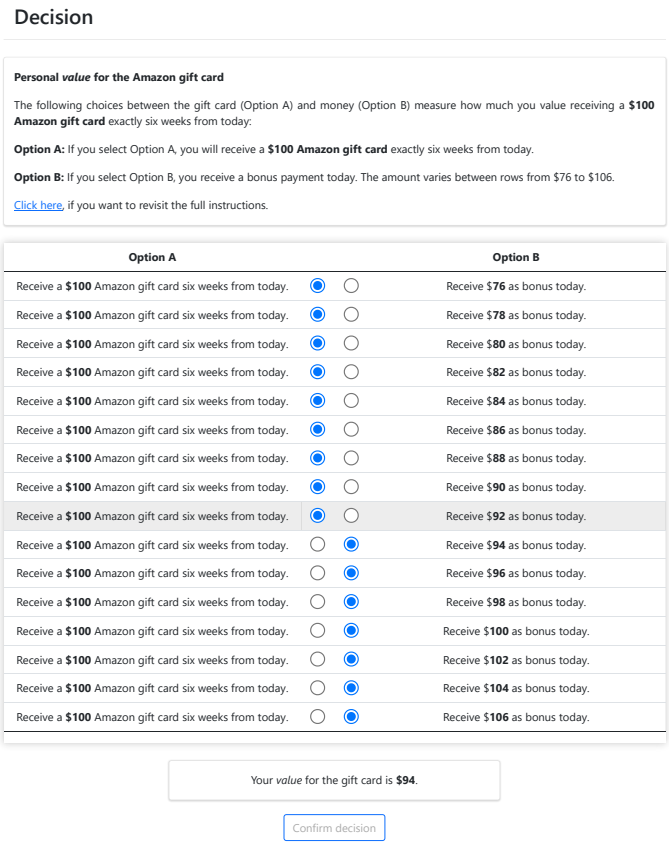


Figure I.101: Giving versus Taking paradigm screen 6

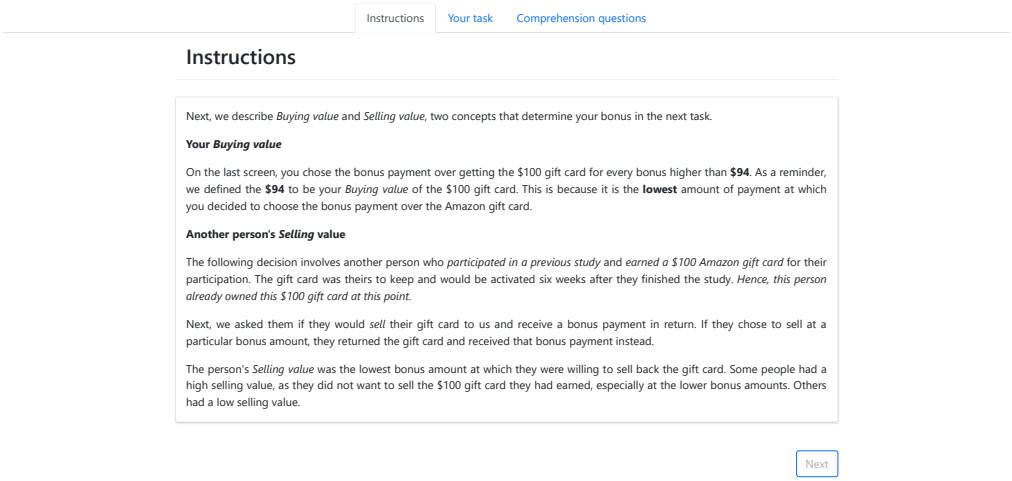


Figure I.102: Giving versus Taking paradigm screen 7

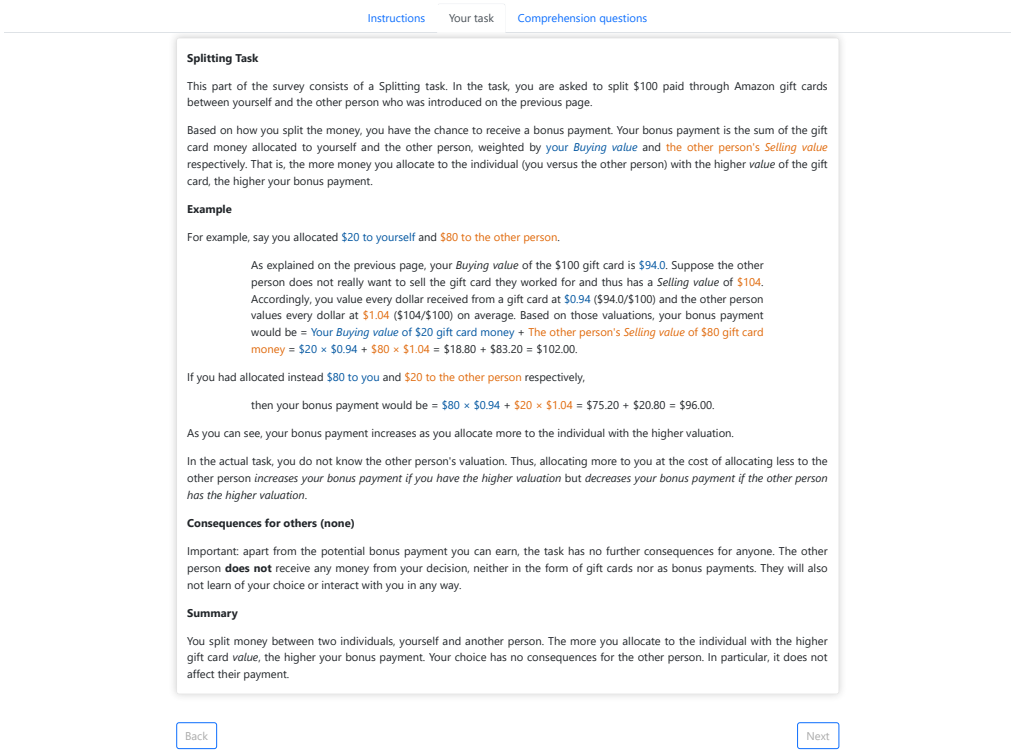


Figure I.103: Giving versus Taking paradigm screen 8

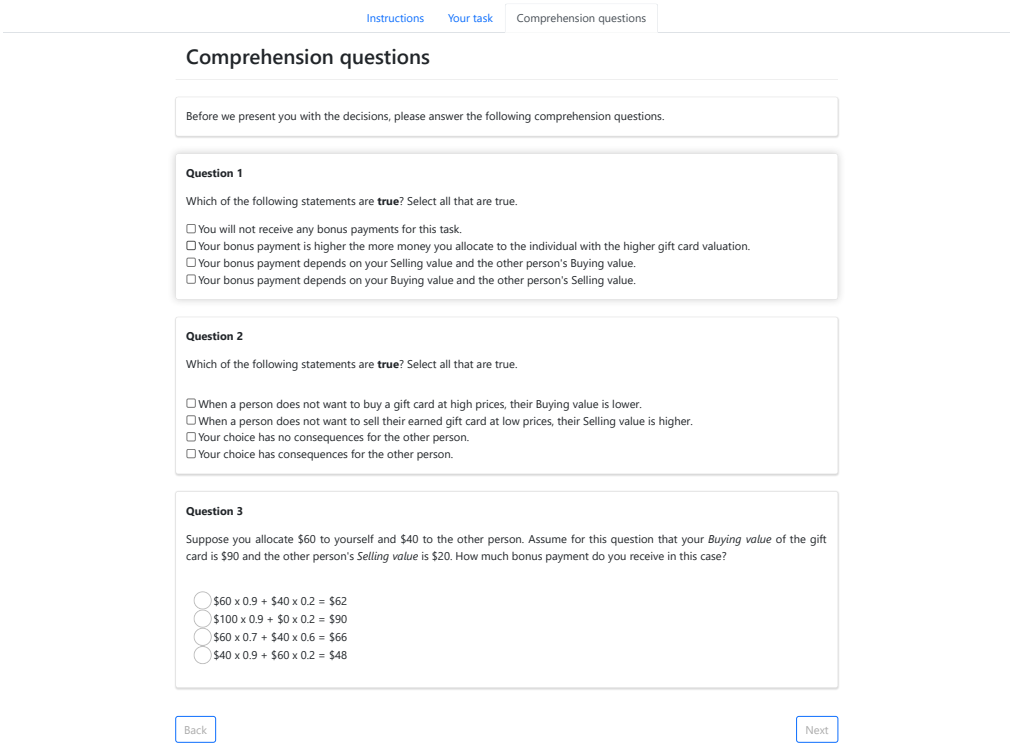


Figure I.104: Giving versus Taking paradigm screen 9

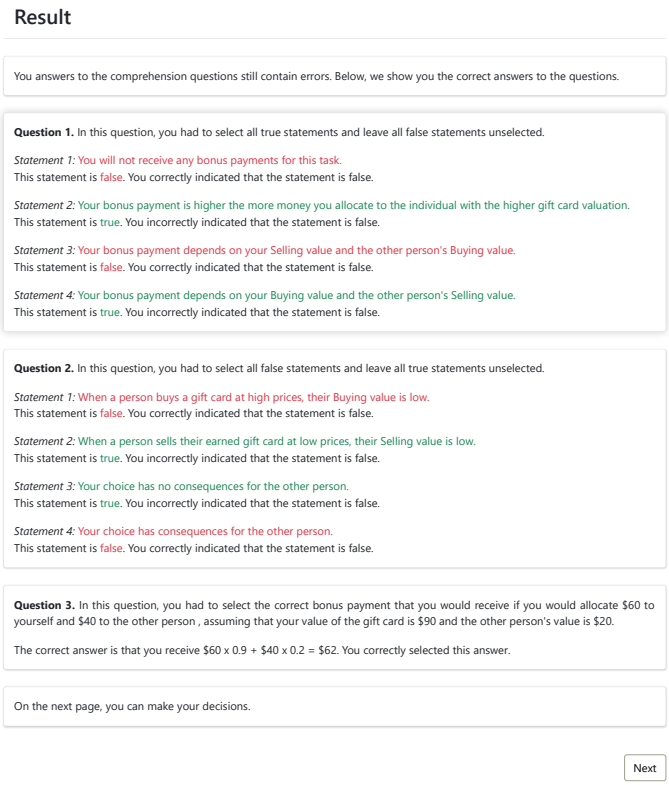


Figure I.105: Giving versus Taking paradigm screen 10

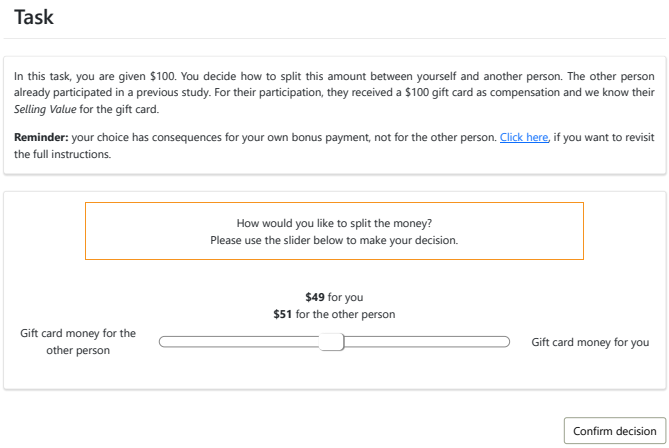
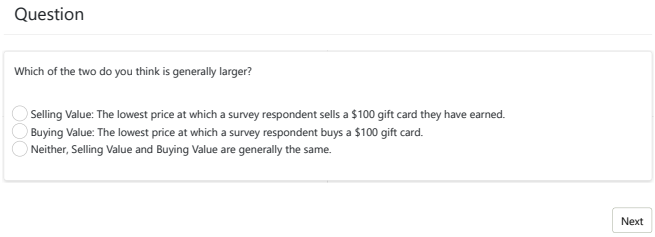


Figure I.106: Giving versus Taking paradigm screen 11



I.4 Belief measurement screens

I.4.1 Ingroup belief measurement screens

Figure I.107: Ingroup belief measurement screen 8

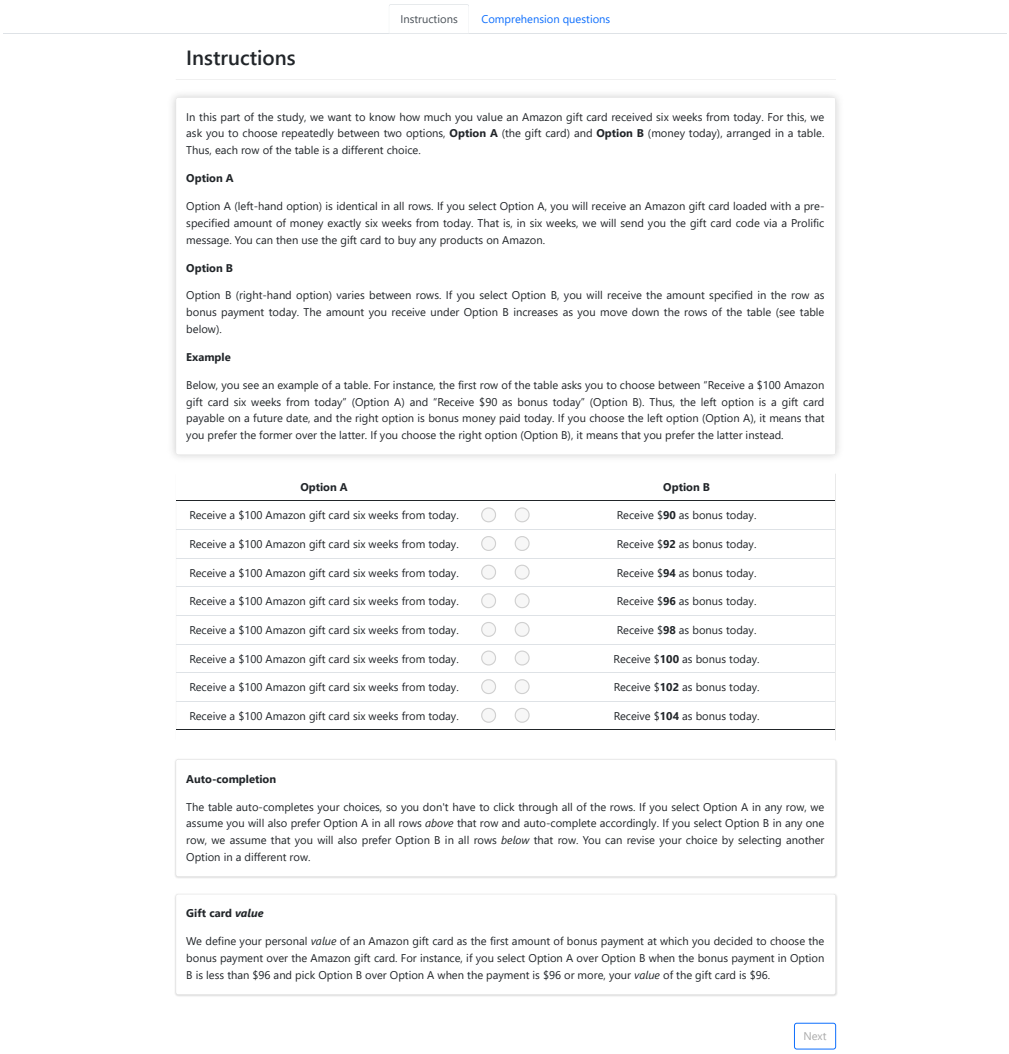


Figure I.108: Ingroup belief measurement screen 9

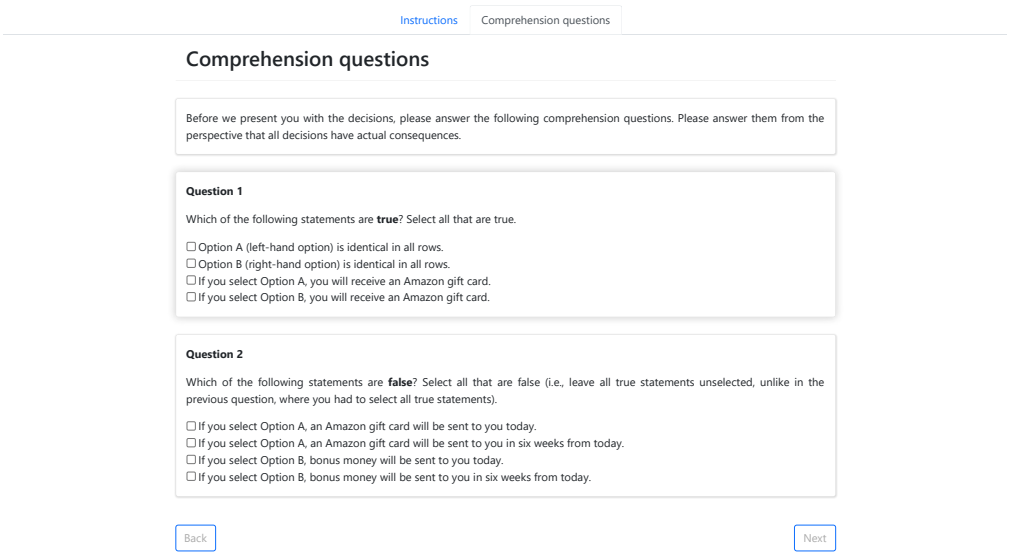


Figure I.109: Ingroup belief measurement screen 10

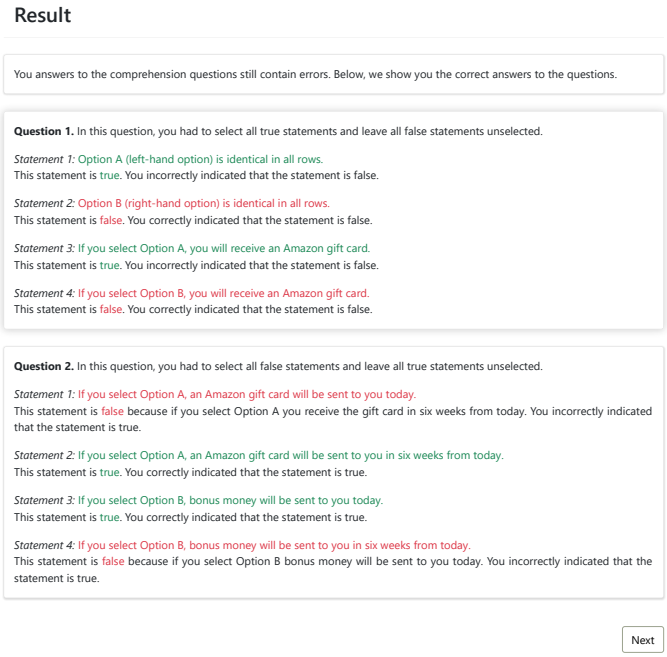


Figure I.110: Ingroup belief measurement screen 1

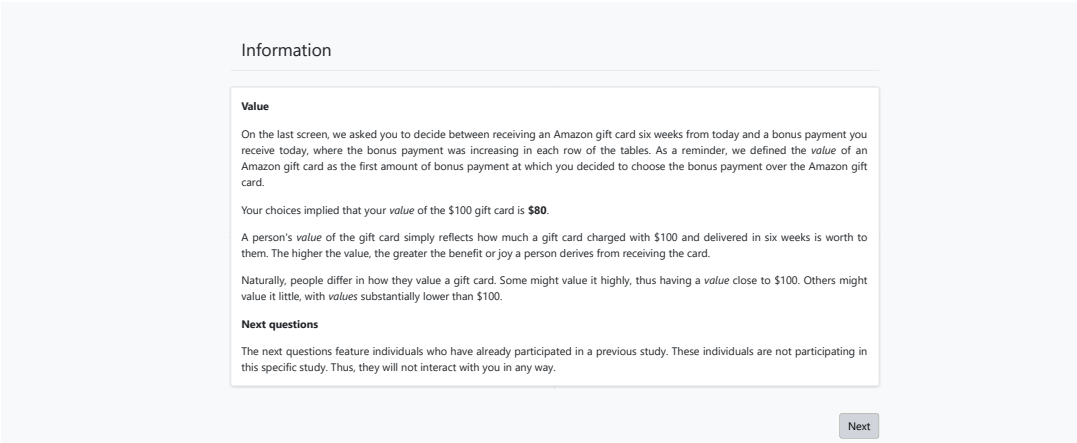


Figure I.111: Ingroup belief measurement screen 2

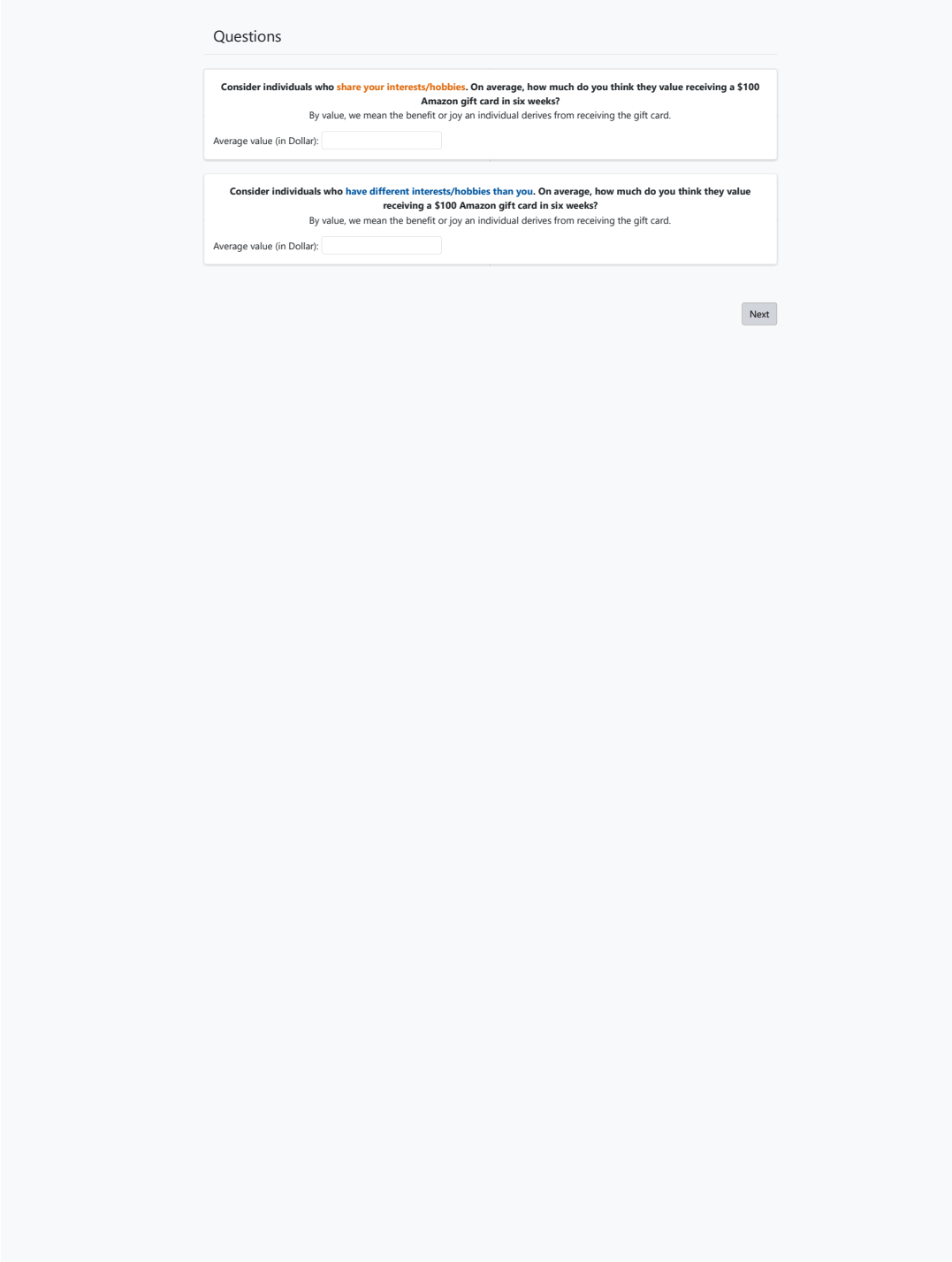


Figure I.112: Ingroup belief measurement screen 3

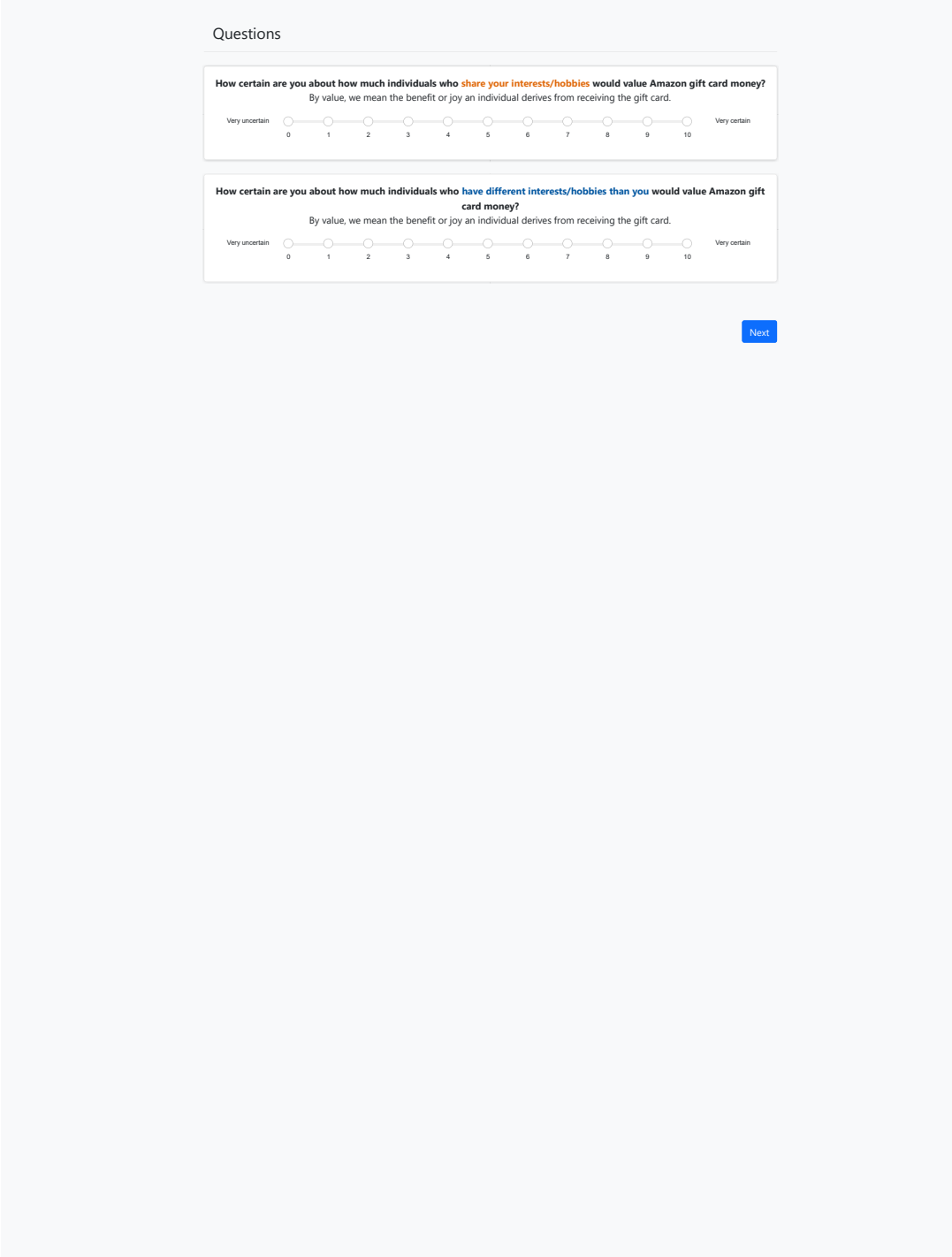


Figure I.113: Ingroup belief measurement screen 4

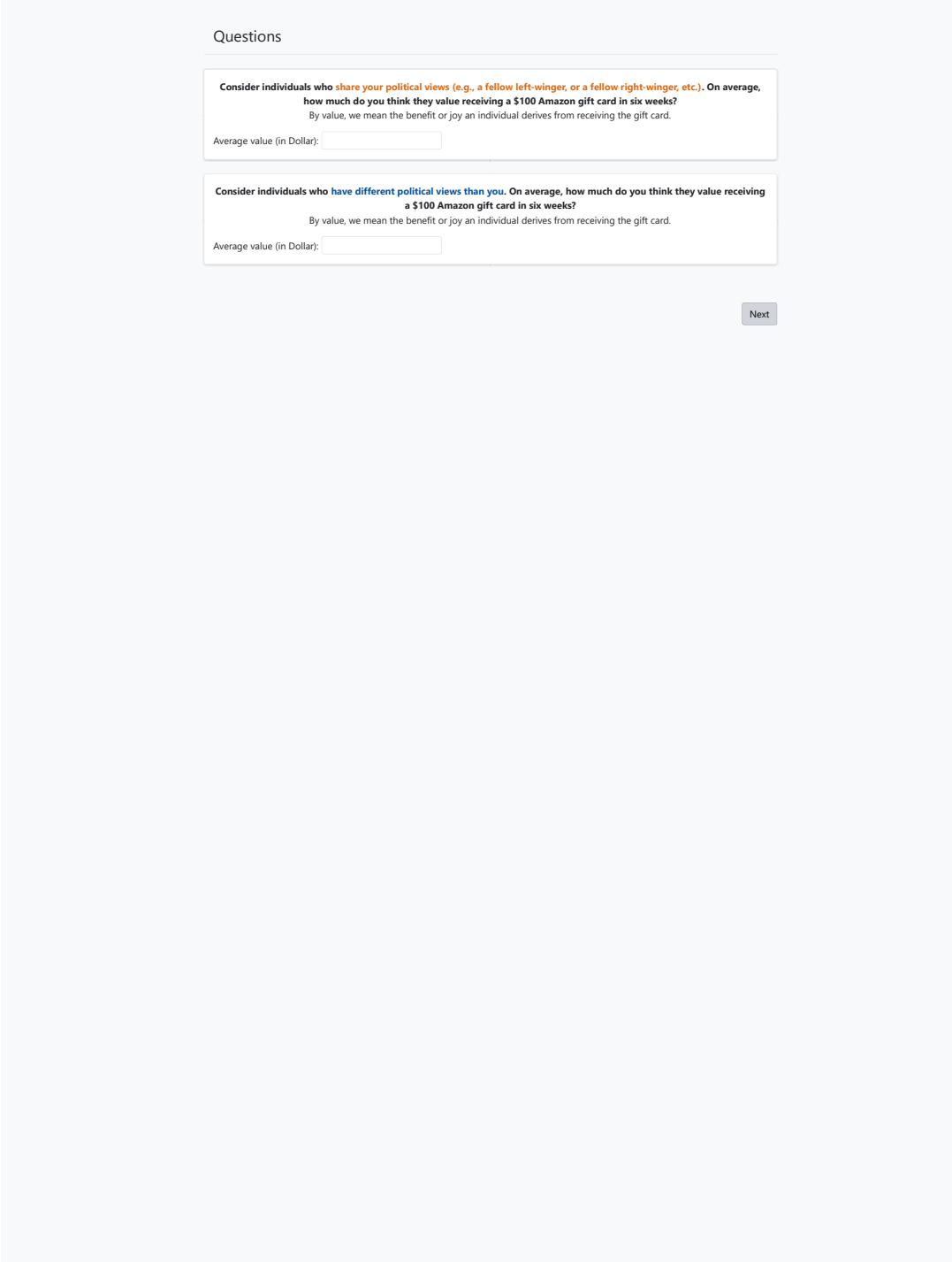


Figure I.114: Ingroup belief measurement screen 5

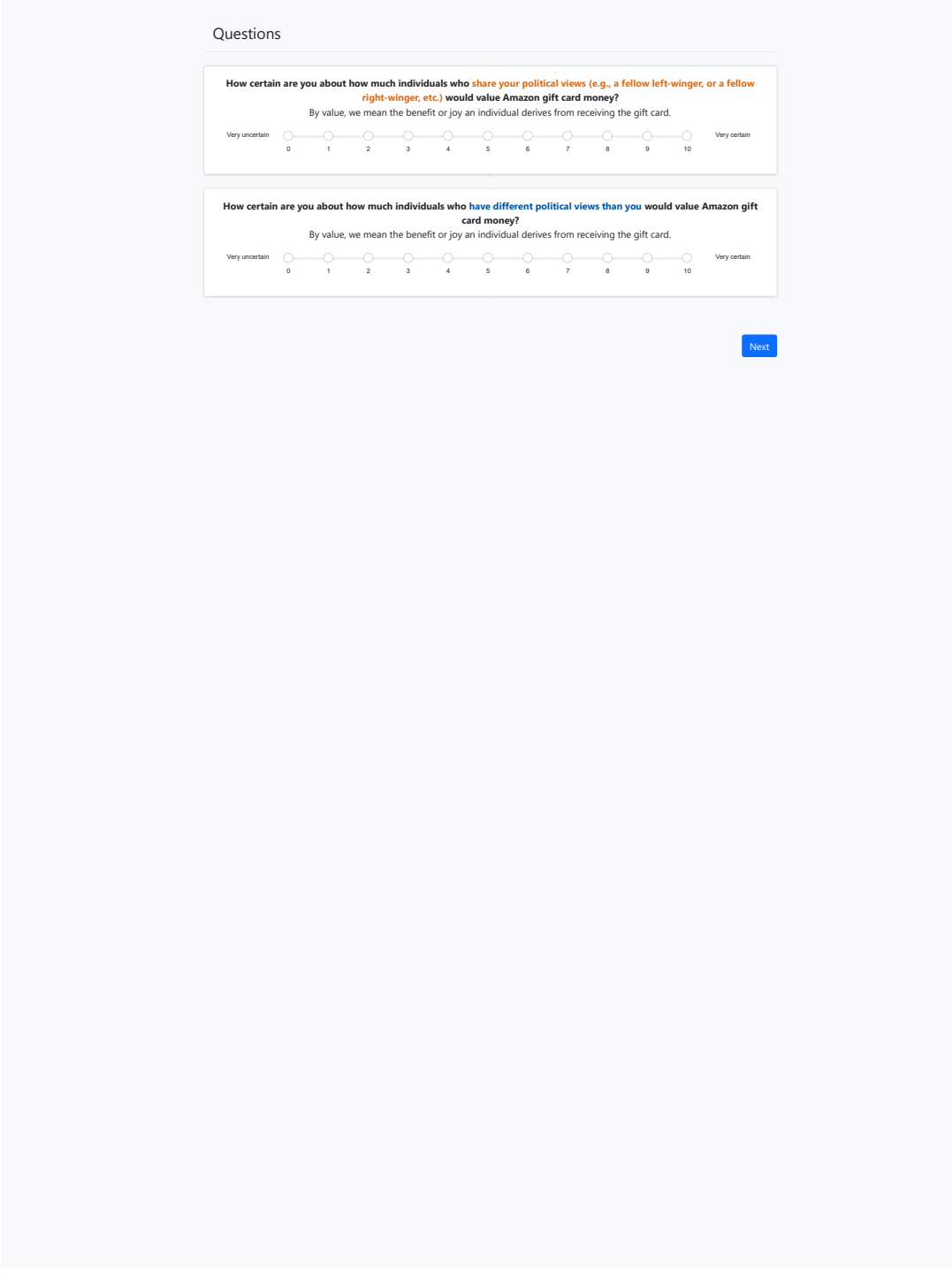


Figure I.115: Ingroup belief measurement screen 6

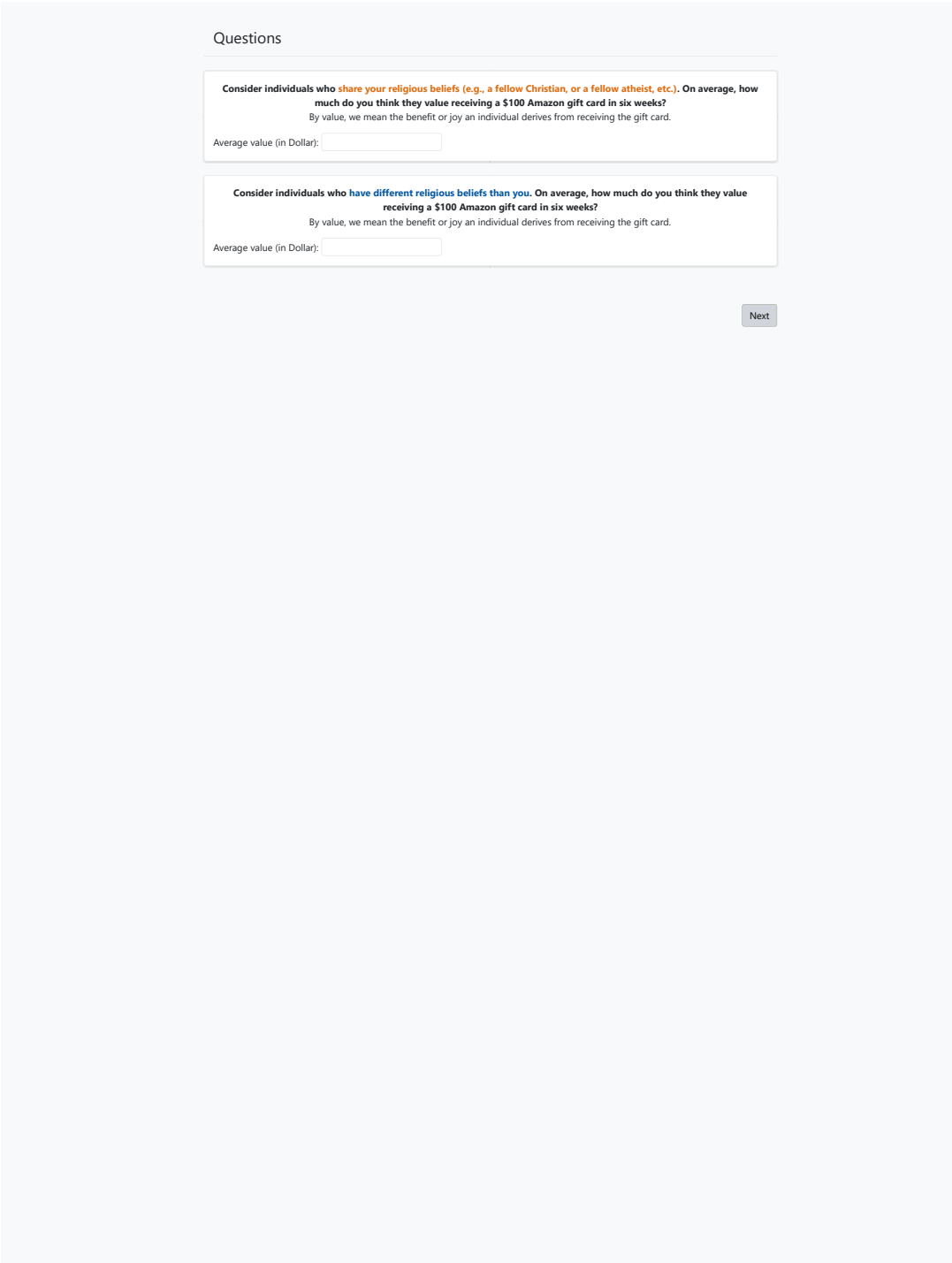


Figure I.116: Ingroup belief measurement screen 7

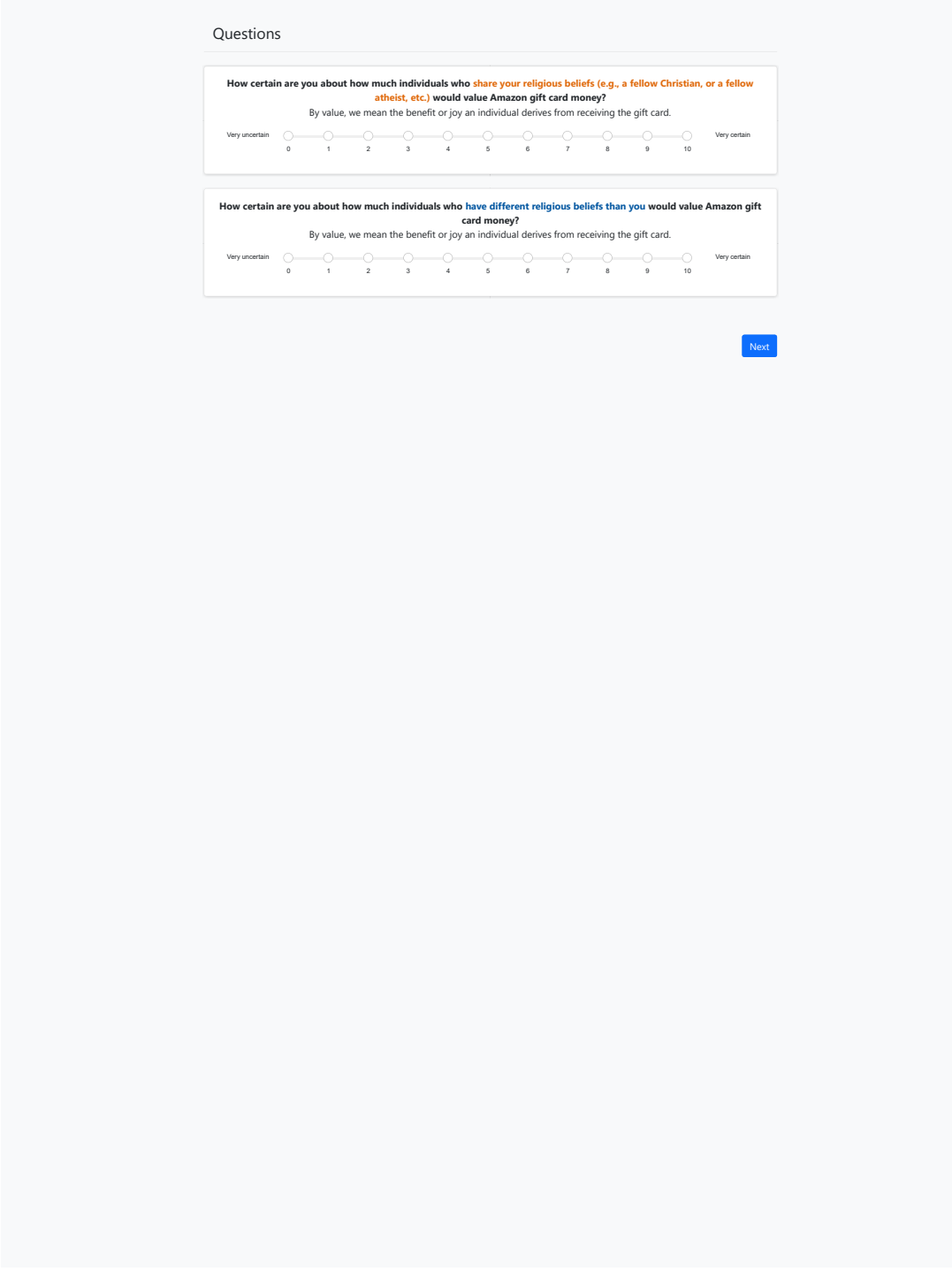


Figure I.117: Ingroup belief measurement screen 8

I.4.2 Self/other belief measurement screens

Figure I.118: Self/other belief measurement screen 8

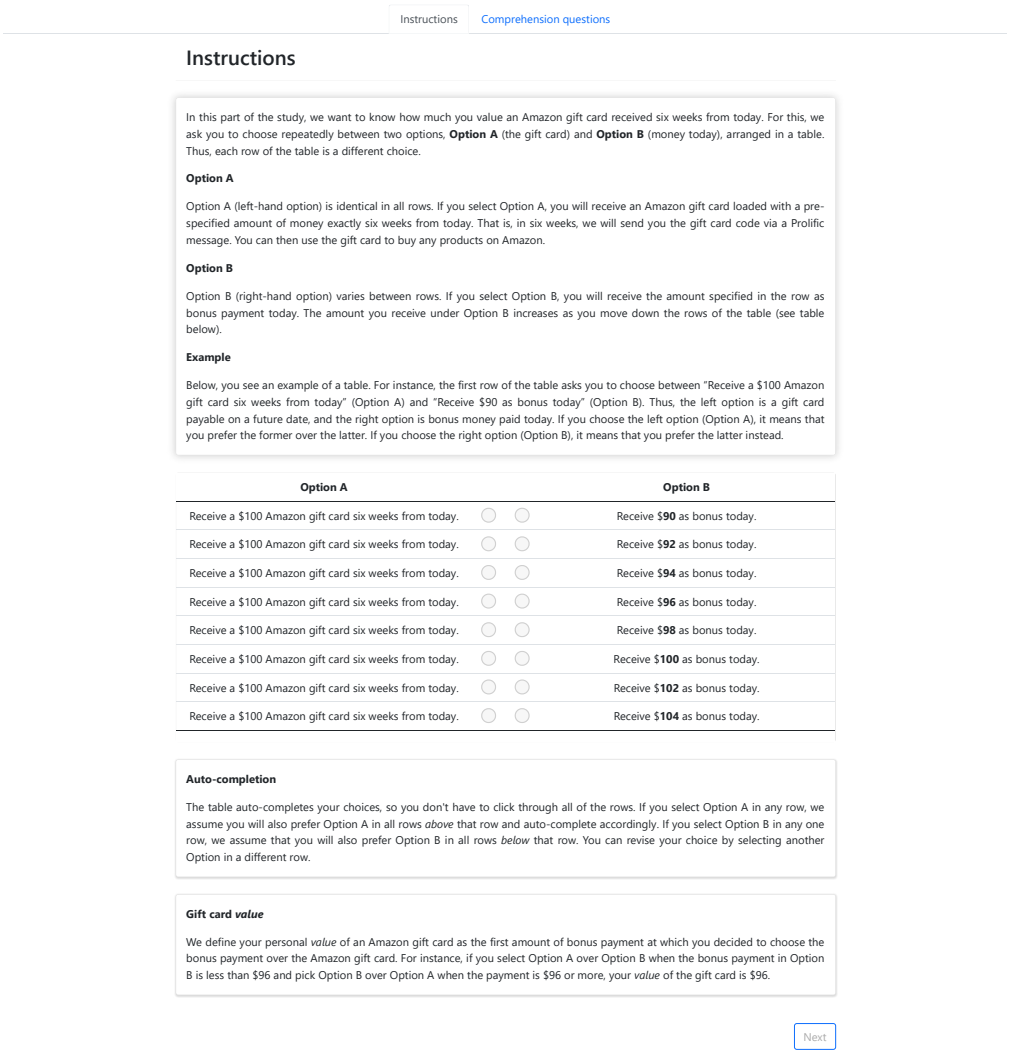


Figure I.119: Self/other belief measurement screen 9

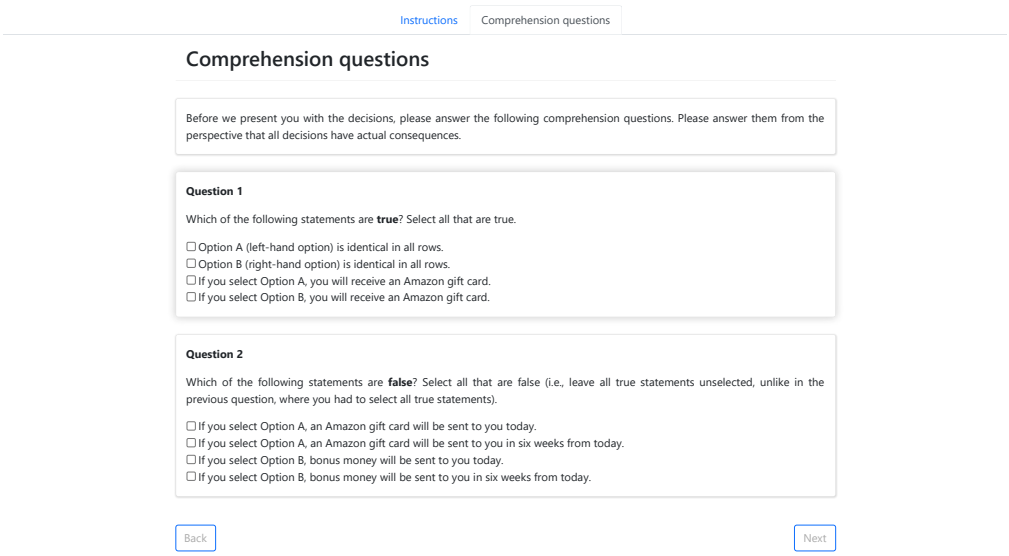


Figure I.120: Self/other belief measurement screen 10

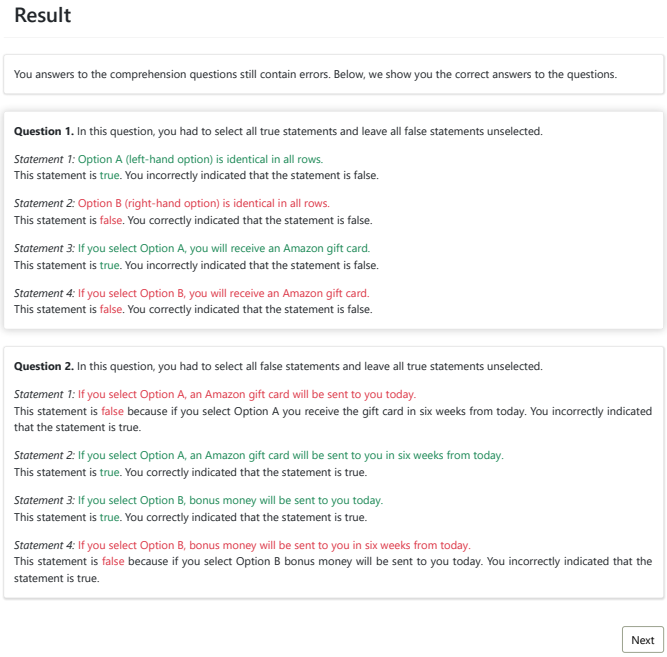


Figure I.121: Self/other belief measurement 1

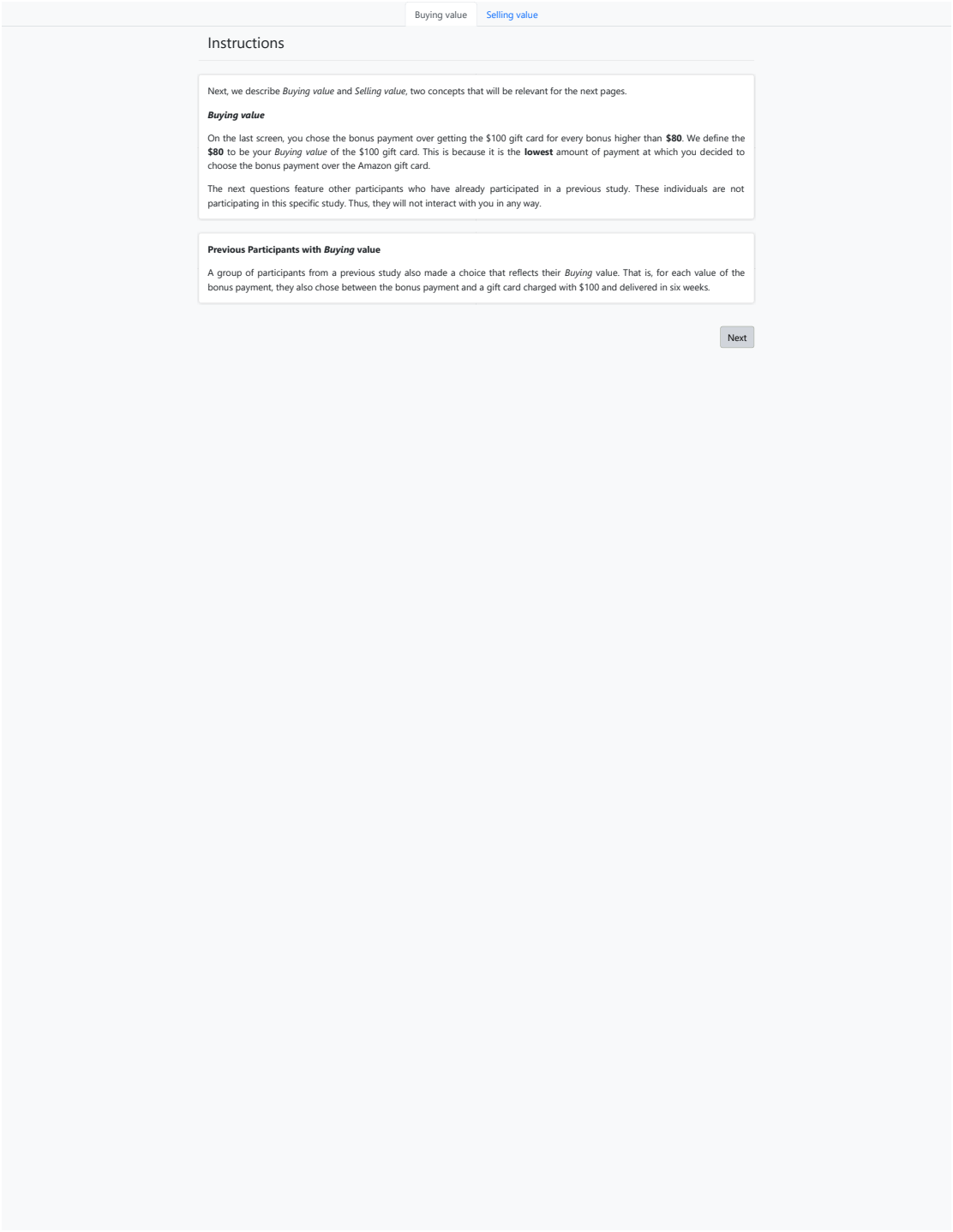


Figure I.122: Self/other belief measurement 2

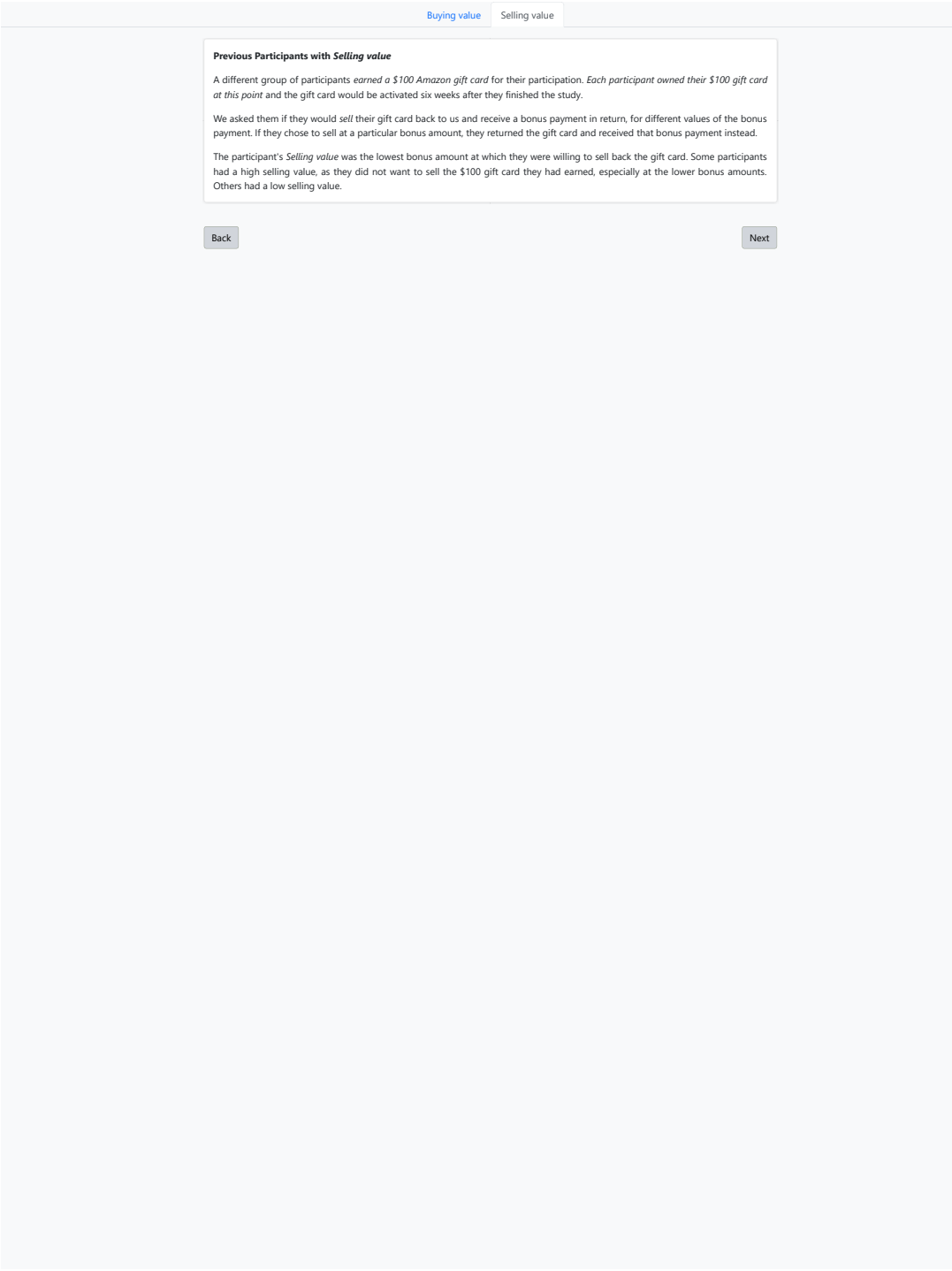


Figure I.123: Self/other belief measurement 3

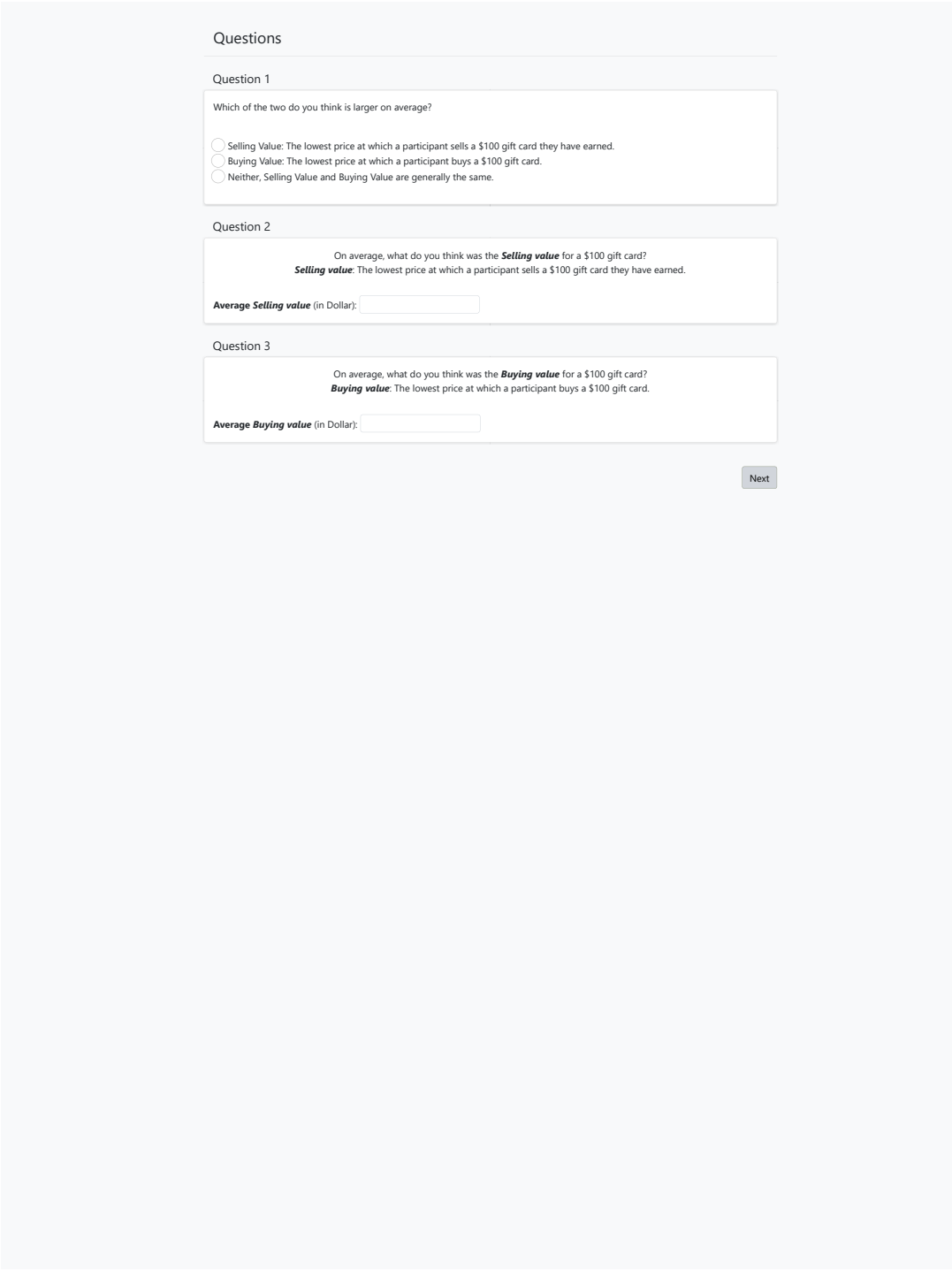


Figure I.124: Self/other belief measurement 4

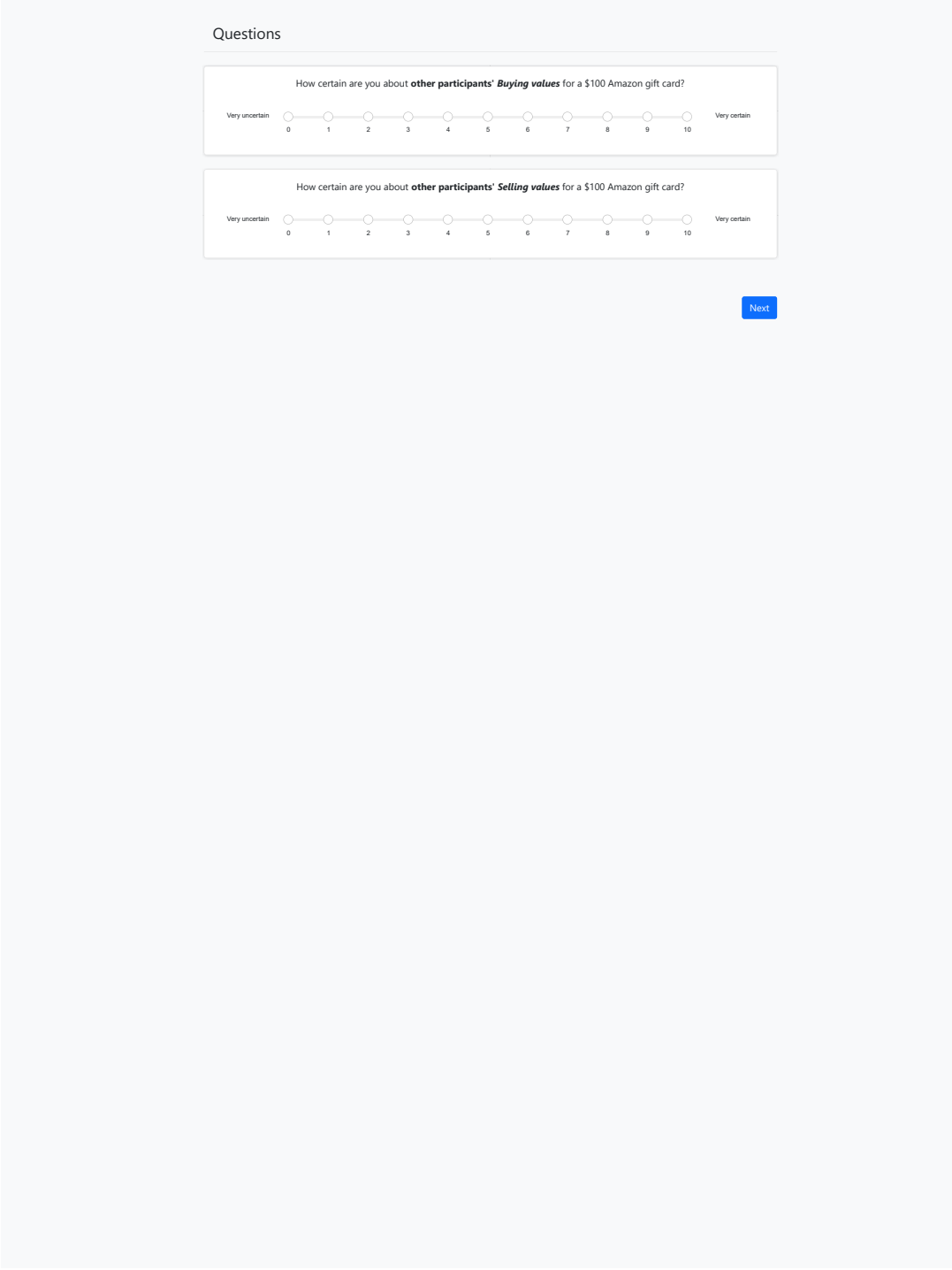


Figure I.125: Self/other belief measurement 5

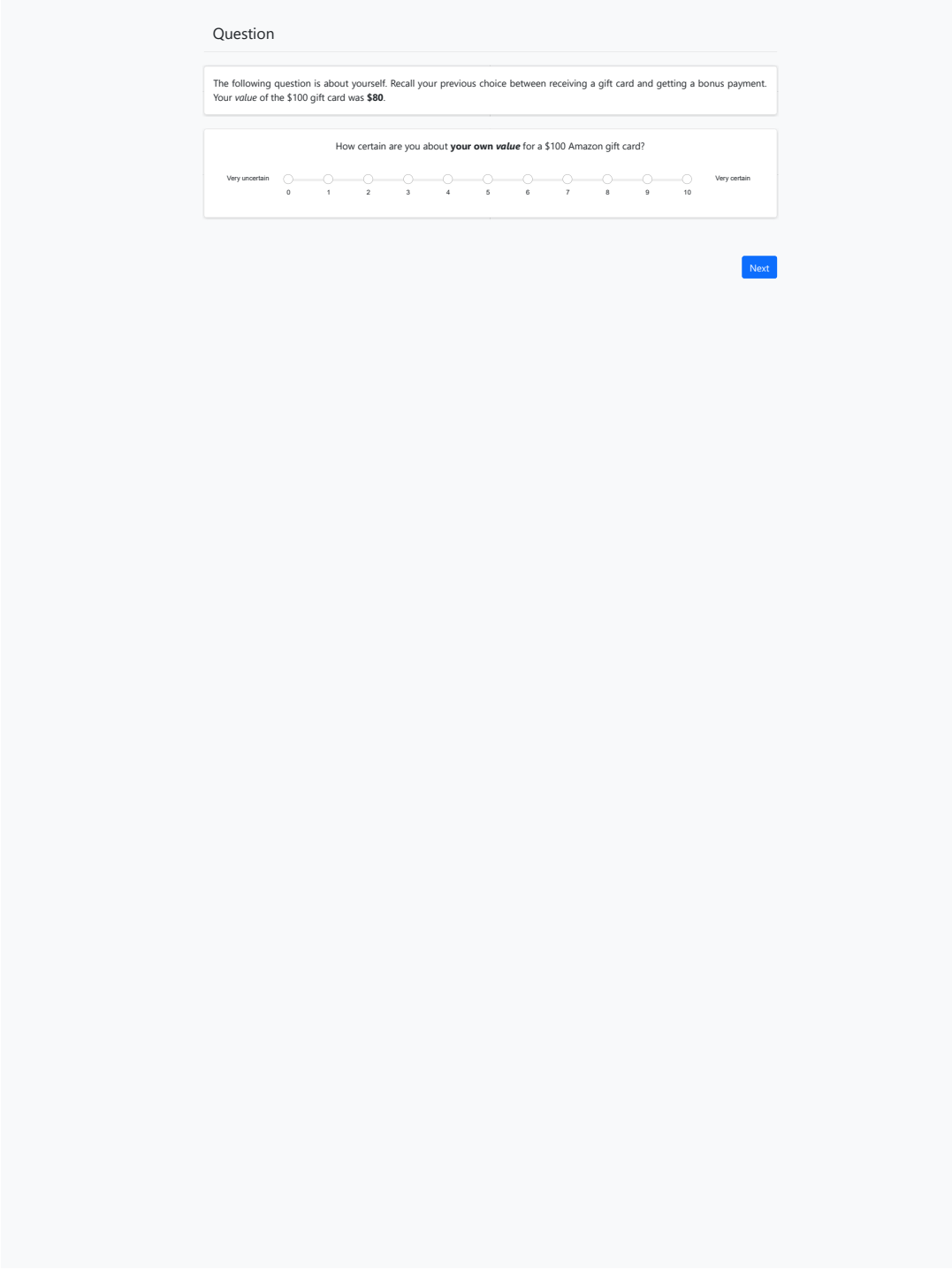


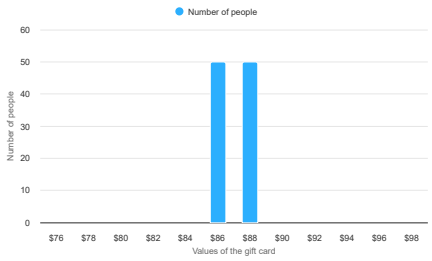
Figure I.126: Self/other belief measurement 6

Questions

In the following, you face two questions. In each, we first show you a figure. Then, we ask you a question about the figure.

In each figure, we plot how a group of 100 people *value* the Amazon gift card. On the horizontal axis, you see the values. The values are even numbers between $76 and $98. The vertical axis shows the height of the bar which equals the number of people having the respective value.

Figure for Question 1



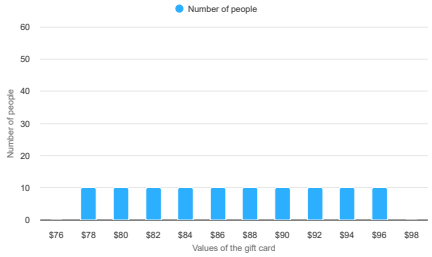
Question 1

In this figure, 50 people have a *value* of $86 and 50 people have a *value* of $88. Suppose we randomly pick one of the 100 people from this group.

How certain are you about how much the randomly chosen person would value the Amazon gift card money?

Very uncertain 0 1 2 3 4 5 6 7 8 9 10 Very certain

Figure for Question 2



Question 2

In this figure, there are 10 people at each of the 10 *values* between $78 and $96. Suppose we randomly pick one of the 100 people from this group.

How certain are you about how much the randomly chosen person would value the Amazon gift card money?

Very uncertain 0 1 2 3 4 5 6 7 8 9 10 Very certain

Next